PB81-136103

The SEAS (Strategic Environmental Assessment System) Region Model An Assessment of Current Status and Prospects

Johns Hopkins Univ.
Baltimore, MD

Prepared for

Environmental Protection Agency
Washington, DC

May 79

U.S. Department of Commerce
National Technical Information Service
NTIS

TECHNICAL REPORT DATA

(Please read Instructions on the reverse before completing)

1. REPORT NO. EPA-600/7-79-119	2.	3. RECIPIENT'S ACCESSION NO. PB81 136103
4. TITLE AND SUBTITLE The SEAS Region Model: An Assessment of Current Status and Prospects		5. REPORT DATE May 1979 issuing date
		6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) T. R. Lakshmanan, Peter Kroll, Bill Barron, Lata Chatterjee, M. Pappas		8. PERFORMING ORGANIZATION REPORT NO.
9. PERFORMING ORGANIZATION NAME AND ADDRESS The Johns Hopkins University Baltimore, Maryland 21218		10. PROGRAM ELEMENT NO. 1NE 827C
		11. CONTRACT/GRANT NO. Grant No: R804869-01
12. SPONSORING AGENCY NAME AND ADDRESS U.S. Environmental Protection Agency Office of Research and Development Office of Energy, Minerals & Industry (RD-681) Washington, D.C. 20460		13. TYPE OF REPORT AND PERIOD COVERED
		14. SPONSORING AGENCY CODE EPA/600/7

15. SUPPLEMENTARY NOTES
This project is part of the EPA-planned and coordinated Federal Interagency Energy/Environment R&D Program.

16. ABSTRACT

This document describes the research performed in response to a recognized need for improvement of the regional component (REGION) of the U.S. Environmental Protection Agency's Strategic Environmental Assessment System (SEAS), a computer model of the nation's economic-environmental-energy inter-actions designed to assess alternative environmental and energy policy scenarios.

The research reported here consists of two parts. Part I is an assessment of the quality of the regional economic data base developed by the REGION module of SEAS. A comparison of this data base with regional output data computed by U.S. Census Bureau leads to a recommendation that SEAS base year data base be recalibrated according to the Census data, taking into account spatial variations in sector output per worker.

Part II examines alternative regional projection methods for possible inclusions in SEAS. The method of location quotients to derive regional input-output tables from National I-O is re-evaluated with encouraging results. Reviews of alternative existing models were then performed, with their handicaps against inclusion in SEAS discussed. A new Multiregional Model of Regional Economy and Energy Demand (MREED) is developed as an alternative. Finally, a model to assess regional impacts of changes in a region's energy prices or final demand is presented.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
Energy Policy	Integrated Assessment	97G

18. DISTRIBUTION STATEMENT RELEASE TO PUBLIC	19. SECURITY CLASS *(This Report)* Unclassified	21. NO. OF PAGES 162
	20. SECURITY CLASS *(This page)* Unclassified	22. PRICE

EPA Form 2220-1 (9-73)

United States
Environmental Protection
Agency

Office of Energy, Minerals, and
Industry
Washington DC 20460

EPA-600/7-79-119
May 1979

PB81-136103

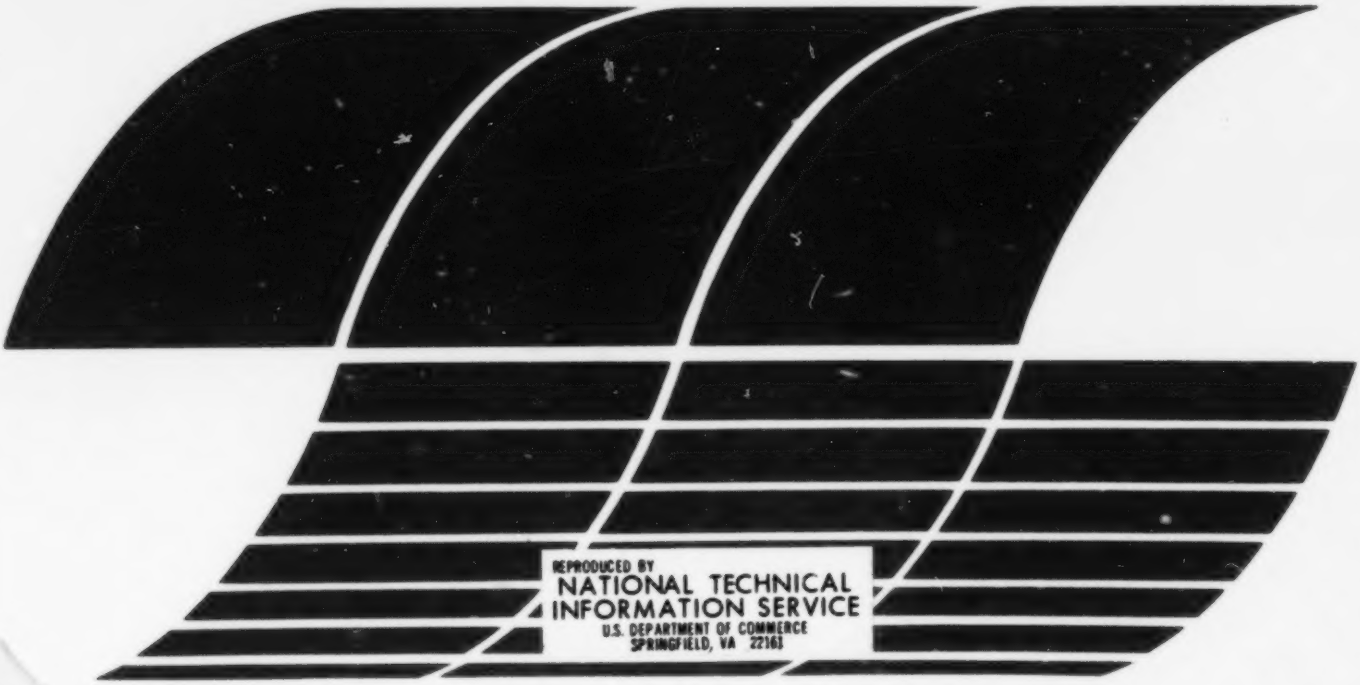
Research and Development

EPA

The SEAS Region Model

An Assessment of Current Status and Prospects

Interagency Energy/Environment R&D Program Report



REPRODUCED BY
NATIONAL TECHNICAL
INFORMATION SERVICE
U.S. DEPARTMENT OF COMMERCE
SPRINGFIELD, VA 22161

RESEARCH REPORTING SERIES

Research reports of the Office of Research and Development, U.S. Environmental Protection Agency, have been grouped into nine series. These nine broad categories were established to facilitate further development and application of environmental technology. Elimination of traditional grouping was consciously planned to foster technology transfer and a maximum interface in related fields. The nine series are:

1. Environmental Health Effects Research
2. Environmental Protection Technology
3. Ecological Research
4. Environmental Monitoring
5. Socioeconomic Environmental Studies
6. Scientific and Technical Assessment Reports (STAR)
7. Interagency Energy-Environment Research and Development
8. "Special" Reports
9. Miscellaneous Reports

This report has been assigned to the INTERAGENCY ENERGY-ENVIRONMENT RESEARCH AND DEVELOPMENT series. Reports in this series result from the effort funded under the 17-agency Federal Energy/Environment Research and Development Program. These studies relate to EPA's mission to protect the public health and welfare from adverse effects of pollutants associated with energy systems. The goal of the Program is to assure the rapid development of domestic energy supplies in an environmentally-compatible manner by providing the necessary environmental data and control technology. Investigations include analyses of the transport of energy-related pollutants and their health and ecological effects; assessments of, and development of, control technologies for energy systems; and integrated assessments of a wide range of energy-related environmental issues.

This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161

EPA-600/7-79-119
MAY 1979

THE SEAS REGION MODEL:

An Assessment of Current Status and Prospects

by

T. R. Lakshmanan*
Peter Kroll+
Bill Barron+
Lata Chatterjee*
M. Pappas+

*Boston University
+The Johns Hopkins University

EPA Grant R8048690-01

Project Officer

Richard Ball, Ph.D.
Environmental Protection Agency

Prepared for
Office of Research and Development
U.S. Environmental Protection Agency
Washington, D.C. 20460

DISCLAIMER

This report has been reviewed by the Office of Energy, Minerals, and Industry, U.S. Environmental Protection Agency, and approved for publication. Approval does not signify that the contents necessarily reflect the views and policies of the U.S. Environmental Protection Agency, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

ABSTRACT

This document describes the research performed in response to a recognized need for improvement of the regional component (REGION) of the U.S. Environmental Protection Agency's Strategic Environmental Assessment System (SEAS), a computer model of the nation's economic-environmental-energy interactions designed to assess alternative environmental and energy policy scenarios.

The research reported here consists of two parts. Part I is an assessment of the quality of the regional economic data base developed by the REGION module of SEAS. A comparison of this data base with regional output data computed by U.S. Census Bureau leads to a recommendation that SEAS base year data base be recalibrated according to the Census data, taking into account spatial variations in sector output worker.

Part II examines alternative regional projection methods for possible inclusions in SEAS. The method of location quotients to derive regional input-output tables from National I-O is re-evaluated with encouraging results. Reviews of alternative existing models were then performed, with their handicaps against inclusion in SEAS discussed. A new Multiregional Model of Regional Economy and Energy Demand (MREED) is developed as an alternative. Finally, a model to assess regional impacts of changes in a region's energy prices or final demand is presented.

This report was submitted in fulfillment of Grant No. R8048690-01 by The Johns Hopkins University under the sponsorship of the U.S. Environmental Protection Agency. This report covers the period from October 1976 to February 1978 and work was completed as of February 1978.

CONTENTS

Page

Abstract.. iii

Figures..vi

Tables...vii

Acknowledgment.......................................viii

1. Overview and Summary.............................1

2. An Assessment of the Data Base of REGION.........19

3. An Assessment of Location Quotient Methods of Regionalizing the National Input-Output Model.......................................50

4. A Review of Alternative Regional and Interregional Models.........................95

5. A Multi-Regional Model of Regional Economy and Energy Demand (MREED)...................116

6. A Model to Assess Regional Economy Effects of Changes in Final Demand on Energy Prices......................................130

7. Appendix -- Data Sources for Chapter 3.........140

References...146

Preceding page blank

FIGURES

Number Page

1-1 The SEAS System....................................3
4-1 Flowchart of READ Model.........................110
4-2 Energy Demand in READ...........................111
5-1 Structure of MREED Model........................120
5-2 The MREED Model.................................121

TABLES

Number Page

1-1 SEAS/CM Functions for Selected States............7
2-1 Purification of Census of Manufactures Values of Shipments..........................25
2-2 1972 Price Indices for INFORUM Sectors...........26
2-3 Selected National Sector Outputs for 1972........27
2-4 SEAS/CM Fractions for Selected States............28
2-5 Effect of O/W Modifier on SEAS/CM Fractions for Selected States and Selected Sectors.......31
2-6 - 2-39 Detailed Tables for each of the 34 Sectors.......33
3-1 1967 Washington 50-Sector Input-Output Table.....60
3-2 Employment, Output, and Location Quotients for Washington and the U.S....................66
3-3 Comparison Between Direct Coefficients and Tables Obtained from Survey and Short-Cut Methods.......................................78
3-4 Comparison Between $(I-A)^{-1}$ Matrices Obtained From Survey and Location Quotient Methods......80
3-5 Comparison Between Actual and Estimated Output Computed with Final Demand as Given...........81
3-6 Comparison Between Intermediate Outputs Obtained From Survey and Short-Cut Methods.....83
3-7 Comparison Between Exports and Imports Obtained From Survey and Short-Cut Methods.....85
3-8 Comparison Between Tables Obtained From Survey and Balancing Procedure.................87
5-1 Sectors Proposed for MREED Phase I.............119
7-1 Sources Used for Employment Data for Industry Sectors............................140
7-2 Sources Used for Employment Data, by Industry Sector.............................141

ACKNOWLEDGEMENTS

The authors would like to take this opportunity to acknowledge the invaluable help received during the conduct of the research reported here from many individuals too numerous to mention. Professor Anne Carter of Brandeis University provided invaluable help throughout the project reacting to and refining many ideas presented to her in her own inimitable style. This study owes a deep debt to her. We are also thankful to Professor William Miernyk of West Virginia University, who reviewed our efforts at assessing the location quotient method of regionalizing the national I-O table.

This study would not have been possible without the encouragement, discussions and reviews provided by some key members of the SEAS model development and policy analysis team. Particular thanks are due in this regard to Dr. Richard Ball, the project monitor at the U.S. Environmental Protection Agency (EPA), Mr. Edward Williams of U.S. Department of Energy and Mr. Sam Ratick, of U.S. EPA.

We are also grateful to anonymous reviewers of our draft report for valuable comments. The responsibility for any remaining errors of omission and commission is entirely ours.

T. R. Lakshmanan

CHAPTER 1

OVERVIEW AND SUMMARY

The last decade has witnessed a widespread concern with the quality of the natural environment -- the quality of air, water and land -- and the adequacy of the natural resources (particularly energy) to sustain economic growth. This concern stems from such factors as the adverse environmental factors of economic growth, the unrelenting growth of energy use, and the increasing participation of the government in decisions pertaining to environmental protection and energy supply.

This concern has expressed itself in the passage and implementation of a variety of environmental protection laws and regulations and energy conservation and supply augmentation measures. Such laws and regulations have far-reaching economic and environmental consequences. For instance, the regulations requiring industries to abate environmental pollution lead to increased costs of production, and in turn to higher prices, falling output in those industries and reduced employment and income in the region where such industries are located. There may be, at the same time, growth in industries that supply pollution abatement equipment and services in those or other regions. Over time, the health and economic benefits of higher environmental quality express themselves in changing patterns of consumption. Thus as environmental standards pertaining to industries, automobiles and municipalities are implemented, the mix of goods and services in the nation and the region changes. Again, it is well known that the magnitude of energy demand and the mix of sources of fuel or energy in the future will greatly affect the quality of land, air and water bodies.

In the context of this intimate association between the economy, environment and energy, a number of public policy questions are posed by recent and proposed environmental and energy development measures. What are the full range of socioeconomic effects of existing and proposed air, water and land pollution abatement provisions? What are the likely effects on the economy of (and the policies to facilitate) the transition from cheap, or abundant energy and from a reliance on oil and gas to other more expensive sources of energy? What is the nature of the trade-off between additional, or lower cost energy and environmental quality?

Considerable analytical effort has been devoted, in the last half decade, to these policy issues by the development and use of large models at the interface between the economy, environment and energy (Hogan and Parikh, 1977; Hitch, 1977; House, 1977; Lakshmanan, 1978). These models, reflecting the vast scope of economic-environmental-energy interactions, are concerned not only with economic issues but with a range of issues of public policy, physical environmental quality processes (e.g. emissions and spatial diffusion of pollutants), flows of energy and alternative technological processes of energy supply. This overwhelming diversity and the intermingling of socioeconomic and physical relationships in economic-environmental-energy modelling has encouraged cooperative endeavors among social scientists, engineers, and physical scientists. The resulting models often represent an integration of the different analytical approaches the different analytical approaches the different disciplines bring.

One of the earliest and more significant examples of such integrated economic-environmental-energy models is the Strategic Environmental Assessment System (SEAS) model. The SEAS model has evolved from a test version implemented in 1973 in the U.S. Environmental Protection Agency (EPA), to a large system of interlinked models of the economy, environment and energy, operationally used for policy analysis at the U.S. EPA and U.S. Department of Energy (DOE) (Lakshmanan and Krishnamoorthi, 1973; House, 1977; Lakshmanan and Ratick, 1977).

The SEAS model has been developed by a large number of individuals drawn from different disciplines -- economics, geography, engineering, systems analysis, policy sciences, and energy technology -- and different institutions -- universities, private research organizations and the U.S. Government.[1] This multidisciplinary, multi-institutional and development environment has made the SEAS model comprehensive in scope and closely linked to policy analysis. SEAS comprises five component subsystems (Figure 1-1):

1. The National Economic Subsystem (consisting of a linked National Input-Output Model (INFORUM), a sector-process disaggregation model and a pollution abatement cost model).

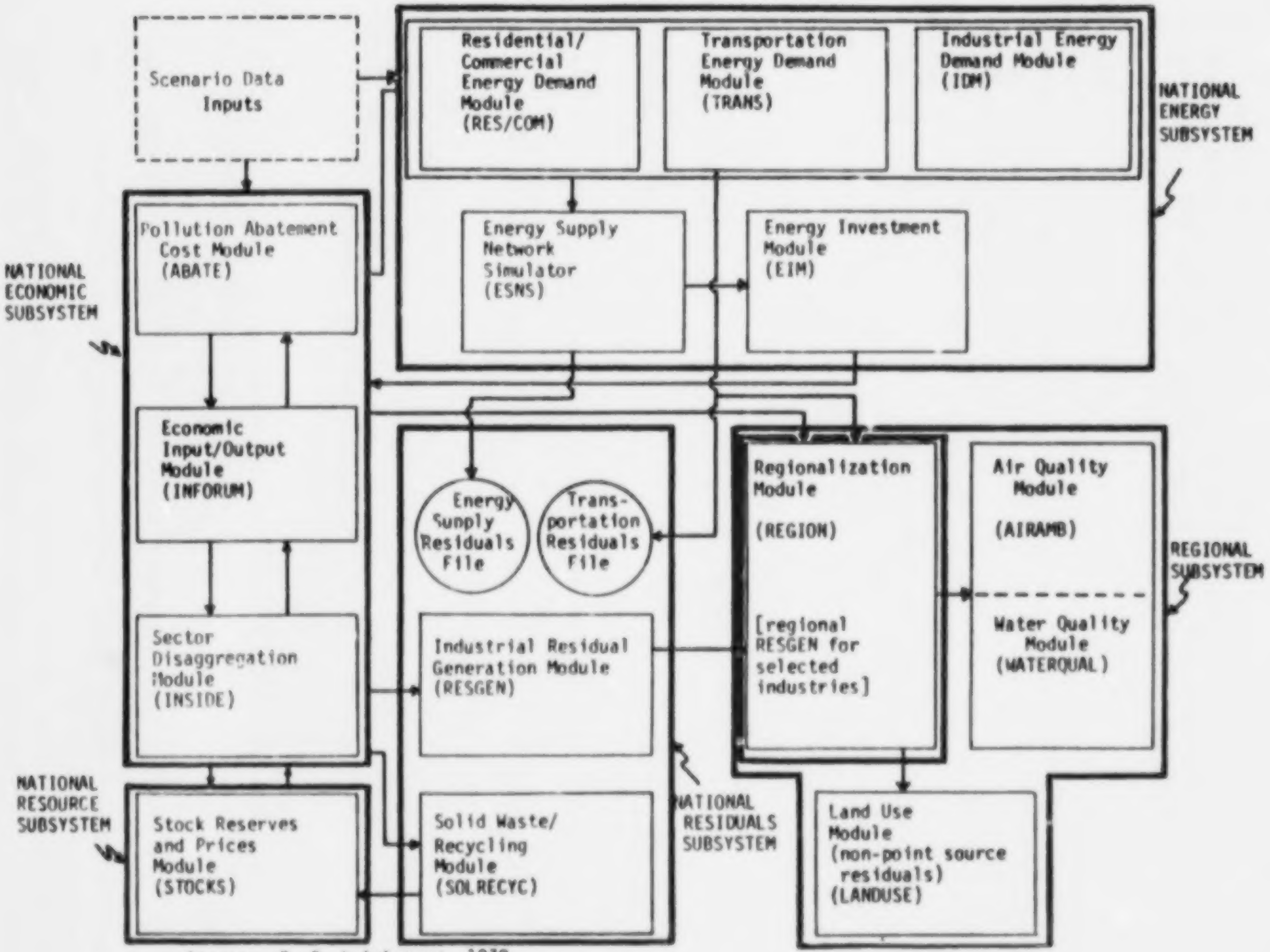
2. The Environmental Subsystem (generating gross and net residuals from production, energy conversion and transportation).

3. Regional Economic and Environmental Subsystems.

4. The Energy Subsystem (generating residential and commercial, industrial and transportation energy demand, total energy resource requirements and energy investment).

---

[1] A selected list of these individuals includes Peter House, T. R. Lakshmanan, Ted Williams, Richard Ball, Roger Shull, Phil Patterson, Sam Ratick, Ron Ridker, William Watson, A. Shapanka, N. Dossani, S. Krishnamoorthi, R. Meyer, R. Doggett, M. Dramer, M. Stern, R. Anderson, B. Wing, E. Lake, R. Ubico, and Peter Kroll.


Scenario Data Inputs
Residential/ Commercial Energy Demand Module (RES/COM)
Transportation Energy Demand Module (TRANS)
Industrial Energy Demand Module (IDM)
NATIONAL ENERGY SUBSYSTEM
Energy Supply Network Simulator (ESNS)
Energy Investment Module (EIM)
NATIONAL ECONOMIC SUBSYSTEM
Pollution Abatement Cost Module (ABATE)
Economic Input/Output Module (INFORUM)
Sector Disaggregation Module (INSIDE)
Energy Supply Residuals File
Trans-portation Residuals File
Industrial Residual Generation Module (RESGEN)
Regionalization Module
(REGION)
[regional RESGEN for selected industries]
Air Quality Module
(AIRAMB)
Water Quality Module (WATERQUAL)
REGIONAL SUBSYSTEM
NATIONAL RESOURCE SUBSYSTEM
Stock Reserves and Prices Module (STOCKS)
Solid Waste/ Recycling Module (SOLRECYC)
NATIONAL RESIDUALS SUBSYSTEM
Land Use Module (non-point source residuals) (LANDUSE)

Source: T. R. Lakshmanan, 1978.

Figure 1-1. The SEAS System

5. National Resource Subsystem (accounting for the effects of depletion of critical raw materials).

From the early steps of its development, this large model has been used for analysis of a wide range of public policies in the U.S. EPA, U.S. DOE, and other agencies. Illustrations of such policies include the economic and environmental effects of resource recovery policies, the annual economic effects of the implementation of the U.S. Clean Air Act and the environmental consequences of the National Energy Plan.

In view of this important role in policy analysis, the SEAS model has been the subject of evaluation. A SEAS Review Panel, chaired by Professors Wassily Leontief and Thomas Crocker, expressed confidence in the overall structure and data bases of SEAS as used at the National level.[2] The panel, however, was less sanguine about the Regional Module (REGION) of SEAS and recommended an improvement of the SEAS Region Model.

In response to this suggestion, a systematic analysis of SEAS Region model structure and data bases was conducted in Johns Hopkins University during 1976-78, under a grant from the U.S. Environmental Protection Agency. The result of this inquiry is detailed in their report.

Two objectives have guided the SEAS Region assessment project. The first objective was to assess the quality of the data bases used in the SEAS Region -- an effort reported in Part I of the document.

The second objective was to recommend an improved structure of the SEAS Region Model. The research work directed at the second objective was organized into tasks. The first task focused on assessment of available alternative models to regionalize national economic data. The second task was the formulation of a regional model designed especially to facilitate environmental and energy policy analysis in the structure and policy analysis context of SEAS. The outcome of the research in both these tasks is the theme of Part II of this report.

We highlight in the remaining portion of this chapter, the salient features of research directed at both of the above objectives. We present the same results in greater detail in the following chapters.

PART I. ASSESSMENT OF THE DATA BASE OF REGION

Introduction

The first task of the research work was to assess the current data base of REGION, the regionalization module of the Strategic Environmental assessment System (SEAS) Model. A brief description of REGION will be presented here. It is followed by a comparison of economic output data for 1972 for selected INFORUM sectors in 15 major states generated by SEAS Region Module against data collected by the U.S. Bureau of the Census for the same year. Discrepancies between the actual state values and the SEAS regional estimates are then interpreted in terms of variations in state labor productivity from the national norm. From the results of this analysis we present some suggestions for future approaches to strengthening REGION's current data base.

---

[2] The Ad Hoc SEAS Review Panel (Co-chairmen: Wassily Leontief and Thomas Crocker), Quality Review of SEAS, (Mimeo), December 3, 1975.

The Regionalization Module of SEAS

REGION allocates national projections generated by other SEAS modules -- economic output (from INFORUM), and pollutant emissions (from RESGEN) -- to counties and reaggregates the estimates to any of several levels of regional classification. No analytical representation of intraregional or interregional relationships of production or consumption is contained in REGION's structure. Rather, national sector output is allocated to regions according to each region's proportional share of national employment (in most cases) in that sector. REGION contains a set of these fractions for each regional classification scheme, for economic output and abatement costs (E-SHARE files) and for residuals generation (P-SHARE). The P-SHARE file, applied to all residual types generated by each industry therefore allocates to a region the same fraction of each pollutant generated by a given national industry.) These base-year (1971) SHARE files are based on county-level employment data and aggregated to the appropriate regional level.

SHARE values for future years are adjusted by indices developed from 1972 OBERS Projections: Regional Economic Activity in the U.S. (Series E Population, BEA/ERS, U.S.G.P.O., 1974). This "OBERS Growth Index" is a ratio of projected earnings of the 2-digit SIC code containing the INFORUM sector to those of the base year (1971). The adjusted SHARE fractions are then normalized to sum to 1.

Several disadvantages are present in sich use of employment as an index in constructing shares for allocation. First, by using employment as a surrogate for output, SEAS incorporates the implicit assumption of constant labor projections themselves employ 'shift-share' methods, extrapolating present trends of industrial location. Third, the pollutants SHARE file assumes that (in most sectors and most regions) the national technology within an industry holds for all regions (although provision exists for the use of other values if data were available). All of these weaknesses raise doubt about the accuracy of the current SEAS regional allocation procedure.

Comparison of Regional Data Bases

SEAS-generated estimates of industrial output at the state level for 1972 were compared with Census Bureau data for that year, presented in the Census of Manufactures. Comparisons of these two data sets were made for 34 selected industries in a sample of 15 large states. However, several preliminary adjustments were first required. The SEAS output (reported in

1971 dollars, hence adjusted for 1972 prices) is categorized by sector according to "purified" sector used in the SEAS INFORUM model. In INFORUM, all output by that industry consists of products of that industry. The Census of Manufactures (CM) data, however, is reported on a state basis only with all products produced by an establishment assigned to the industrial sector manufacturing the product whose output is greatest at that establishment. As a result, the state CM data required "purification" by multiplication by the national ratio of purified to unpurified output by sector, both of whose values are available at a national level in the Census of Manufactures. Finally as the 1972 SEAS outputs are projections of a base year (1971), the national SEAS and CM output figures cannot be presumed to coincide. The state SEAS figures are simply allocations of national projections, so in order for these allocated shares to be comparable to state CM data, the state SEAS estimates were scaled by the ratio of national CM sector output to national SEAS sector output. In this way, REGION's allocation of the correct national totals could be tested.

Once these adjustments were made, the state CM and estimated SEAS values were compared, with substantial variation observed. Measures of these variations are provided in Table 1-1 and detailed tables in Chapter 2. From these results, it is clear that SEAS' allocation of national output to regions involves substantial inaccuracy.

Accounting for Output Per Worker Variation

We had suggested earlier that the SEAS REGION Model, by allocating national output to region using employment shares, ignores regional variations of output per worker in an industry. The consequences of this procedure were explored next.

The ratio of state to national output per worker by industry (obtained from the Census of Manufactures) was used as an adjustment factor applied to SEAS state estimates. This adjustment step generally caused the SEAS REGION estimates to move closer to Census data, but frequently the full difference was not accounted for. If it could be incorporated as a second factor in REGION's shares, however, substantial improvement may be obtained, at least for those sectors for which regional data are available. This suggestion is made under the presumption that regional output per worker differentials do not vary greatly in the four or five year intercensal period.

Implications

Considerable differences between SEAS estimates and Census of Manufactures result from the current operations of SEAS' Regionalization module. Output per worker variations among regions may account for some discrepancy, and could be partially accounted for in the REGION module by adjusting the SHARE factors where data are available, assuming regional stability of output per worker relationships. Additional error may be attributable to noise derived from basing the shares on county-level employment, aggregated to the appropriate regional configuration. The

TABLE 1-1. SEAS/CM FRACTIONS FOR SELECTED STATES[a]

SEC-TOR	NY	OHIO	ILL	CAL	TEX	SEC-TOR	NY	OHIO	ILL	CAL	TEX
23	2.03	1.77	1.34	.98	1.26	76	1.00	D	.77	.73	.66[b]
24	.57	1.12	.60	.74	2.49	81	D	1.16	1.23	.85	.98
25	.74	.68	.98	.86	1.23	84	.83	1.67	1.10	D	D
26	.97	.90	1.23	.78	1.38	89	.45	1.27	D	D	D
27	.94	.93	.87	.92	1.45	94	.96	.99	1.20	D	D
31	.74	1.18	1.13	.66	D	95	1.20	.89	1.19	.95	1.14
39	.43	D	1.31	D	D	97	D	.88	1.08	1.51	3.02[b]
40	1.13	.61	1.03	1.20	1.51	98	1.04	.84	1.00	.90	2.41[b]
45	1.42	.79	1.33	.80	1.02	101	D	D	.83	D	D
51	1.17	.98	1.15	.75	.85	103	1.09[b]	1.02	1.16	D	1.77[b]
52	.98	1.07	1.11	.90	1.08	107	1.08	.70	1.03	D	.72[b]
53	.62	1.72	1.09	.61	D	113	.90	D	1.35	D	1.12[b]
54	.79	1.06	.99	1.16	2.11[b]	118	1.03	1.17	.96	D	.69
59	1.21	D	D	D	.61	119	1.56	.95	1.31	1.00	.90
66	.62	1.06	1.32	1.28	1.50[b]	123	D	1.38	1.03	.75	.55
67	1.71	1.29	.77	1.39	1.18	127	1.06	1.13	.98	.99	D
69	D	.61	1.41	.93	1.12	128	.50	.93	1.29	1.16	D

[a] SEAS is modified by the price index and scaling factor; CM is modified by the purification correction factor. No modification has been made for O/W differences.

[b] State accounts for less than 1% of total sector output.

D = Disclosure limitation

flexibility of the current SEAS regionalization procedures allowing various regional classification schemes can be viewed as a trade off against the more tractable alternatives of data collection and modeling on the scale of a larger regional level such as a state or BEA economic area.

These observations lead to the proposition that a regional data base for economic output be calibrated for a base year, say a census year such as 1972. By calibrating to a state level, the best available regional output data can be utilized supplemented by any additional sources where census data are insufficient. By conducting a careful inventory of regional economic data for a base year, the credibility of SEAS' regional output estimates can be enhanced, and its measurement of regional output can be used with greater confidence in regional analysis and by regional users.

PART II. APPROACHES TO IMPROVE SEAS REGION MODEL STRUCTURE

Our efforts at exploring alternatives to the current practice of using OBERS regional projectors in SEAS were structured into three tasks.

First, we explore the potential of extending the input-output (I-O) framework (used at the national level in SEAS) to the regional level. The depiction of a regional economy by an I-O table permits detailed analysis of the structural interrelationships and is useful in a descriptive and predictive sense. However, data collection for such detailed tasks requires expensive surveys and the data is often dated by the time it is collected and used. Hence, various short-cut methods have been proposed to develop regional tables from a given national table, using only a limited amount of regional data. Such methods would be attractive, as they are relatively simple, quick and inexpensive. However, these methods have yet to be proven effective. In this study we test empirically one nonsurvey -- the location quotient -- method of regionalizing the national I-O table to the state level for the State of Washington with encouraging results.

Second, we review various existing regional and interregional models -- econometric and input-output -- for potential adoption with or without modification as the SEAS REGION model. The primary concern of this review was the suitability of these models as tools for analyzing environmental and energy policies at the regional level.

Next, we present a regional model that is specifically designed for environmental and energy policy analysis in the analytical context of SEAS.

Finally, we discuss a model to assess regional economic effects of final demand or energy price changes. We present the results of this study in brief in the rest of the chapter and in detail in Chapters 3 through 6.

A. An Analysis of a Short-Cut Method of Developing Regional Input-Output Models

Introduction --

Various methods have been suggested to derive regional input-output tables without undertaking expensive and time-consuming surveys. The common procedure is to modify a national table using one of a variety of methods. In this study, location quotients were used to attempt to adjust the 1967 U.S. national table[3] to derive the I-O table for the state of Washington for that year.[4]

The Use of Location Quotients to Regionalize National Input-Output Tables --

In regional studies, location quotients have been used as a measure of an industry's representation or importance in a region, relative to its importance in the nation: $LQ_i = (N_i^R/N^R)/(N_i/N)$, where for industry i, N_i^R represents some measure of the industry's importance (e.g., output, value added, or employment) in region R; N^R is total employment in R; N_i is national employment in i; and N is total national employment.

If $LQ_i = 1$, then the industry has the same importance within the region as it does within the nation as a whole. If $LQ_i<1$, its relative importance is less than it is nationally, and if $LQ_i>1$, it is greater.

The application of this simplest location quotient to a national input-output table requires the assumption that regional technology (i.e., input requirements of industries) is similar to that of the nation. Hence, if the region could supply all of regional intermediate demand, the matrix of direct coefficients A would be identical to that of the nation. In the application of location of quotients to national tables, $LQ_i>1$ is interpreted to mean that this situation is indeed the case; industry i is more important within the region than in the nation, and so can provide sufficient supply to all local intermediate demand. Hence, the inputs required by all regional industries from industry i will be met, and national technology (i.e., all A-matrix coefficients A_{ij} in row i) is valid in depicting regional interindustry requirements.

On the other hand, an industry's location quotient $LQ_i<1$ is taken to mean that regional production of i is insufficient to supply regional demand for industry i's goods, and so only the fraction of the national production requirements, $a_{ij} = LQ_iA_{ij}$ for all purchasing sectors j, can be met within the region by industry i. The remainder of the region's requirements must be made up by imports from the nation. To summarize,

---

[3] Social and Economic Statistics Administration, Bureau of Economic Analysis, U.S. Dept. of Commerce, Input-Output Structure of the U.S. Economy: 1967, v. 1-3, Washington, D.C.: U.S. GPO, 1974. (Magnetic Tape obtained from B.E.A.).

[4] W. B. Beyers et al., Input-Output Tables for the Washington Economy, 1967, Seattle, Wash.: Graduate School of Business Administration, University of Washington, 1970.

letting $LQ_i = (N_i^R/N^R)/(N_i/N)$

$$\text{then } a_{ij} = \begin{cases} A_{ij}, \text{ if } LQ_i \geq 1 \\ LQ_i \cdot A_{ij}, \text{ if } LQ_i < 1 \end{cases}$$

where A_{ij} = national direct coefficient from industry i to industry j, a_{ij} = regional direct coefficient, LQ_i = location quotient for producing industry i, N indicates a measure of economic importance, and superscript R indicates the region.

By adjusting the rows of the national direct coefficients matrix in this manner, an estimate of the regional counterpart of the A-matrix table can be derived with a minimum of information, namely only N_i^R and N_i for all industry sectors.

Approach to Assessment of Location Quotient Method --

While methods of regionalizing national input-output tables in general, and location quotient methods in particular, have been previously studied with less than satisfactory results obtained, the potential was seen for areas of further examination and possible improvement.

First, employment has traditionally been used as the location quotient's measure of economic importance, due to the difficulty of obtaining other economic data. In the present study, the use of total gross output figures, an intuitively better measure of economic importance, was compared with the use of industrial employment figures.

Second, the necessity of having both a national table and survey-based regional table (with which to compare the eventual estimate) in compatible forms (in terms of sector definitions, conventions, and aggregation) often required a high degree of aggregation in order to reconcile the sectors of the two tables. We began with a national table in 367-sector detail, and were therefore able to combine sectors to match those of the Washington table as closely as possible while keeping the estimated table at a level of 50 industrial sectors (compared to the original 52 of Washington). While some important discrepancies between the national and Washington conventions could not be removed, it was felt that the definitional closeness, homogeneity of sectors, and sectoring detail of the present study of location quotients provides a substantial improvement of weaknesses of some previous studies. Also, as is evident, a more detailed table is of greater use in regional analysis.

Third, in testing the accuracy of the estimation of the true regional table, no consistent measure has been used by all analysts. As a result, their results are difficult to compare. We decided to use as many of these tests as possible in order to note any improvement attributable to our procedure.

Fourth, in order to gain some idea of the cost-effectiveness of using this short-cut method to obtain a regional input-output matrix, it was decided to test the estimated table against the "naive" alternative of simply using the national table as a model of regional interindustry relationships. Thus, a comparison can be made of three tiers of estimates of the regional table: the naive model using no regional information (NAIVE); one based on employment location quotients (EMPLQ) requiring only data on employment by industry; and an estimate somewhat more demanding of data, one based on output location quotients (OUTPLQ).

Estimation of the Washington State Input-Output Matrix --
As described previously, the 367-sector 1967 national table (Bureau of Economic Analysis) and the survey-based 1967 Washington State table were both aggregated to a compatible scheme of 50 sectors. The table of direct coefficients was then developed from this 50-sector array by dividing each interindustry flow by the purchasing sector's total gross output. At this point, there were then the "true" survey-based Washington table (WASH), and the national table which we also term the "naive" estimate of the Washington economy.

Then, using employment data and output data, respectively, for both the nation and Washington state by industry, location quotients were computed and applied to the national direct coefficients matrix. The two new tables derived, EMPLQ and OUTPLQ, were the basic location quotient-based estimates of the Washington table. For each model, the A-matrix was also used to calculate $(I-A)^{-1}$. (Because the application of location quotients to the national table can never cause the national coefficients to be exceeded, strict multiplication across a row even if $LQ_i > 1$ was examined as an auxiliary test.)

Once A and $(I-A)^{-1}$ were developed for the three models, plus WASH, their ability to predict total gross output with final demand F given was tested, by multiplying $(I-A)^{-1}F = X$, the familiar input-output relationship.

While this derivation and application (and analysis) of location quotient estimates of the regional direct coefficients matrix was the main aim of our study, some additional implications of the estimates were pursued. With each sector's total output known, along with final demand and the intermediate transactions AX, exports can be computed as residuals for each row. These were calculated. In addition, a "balancing" procedure was performed (under two separate conventions) to preserve the consistency of the table, adjusting for exports calculated as negative residuals (and in one case not permitting any exports where $LQ_i < 1$, by strict interpretation of the theory behind location quotients). This balancing caused some adjustments in the A-matrices which were recalibrated.

The results of all these models of the Washington State table were compared with the values of the true (survey-based) Washington table. The results and conclusions are described in the next section.

Analysis of Results --

As previously stated, our main concern in the test of the use of location quotients to regionalize national input-output tables was to accurately derive the matrix of direct coefficients which, given a vector of regional final demands, can be used to calculate total gross output by industry: $(I-A)^{-1}F = X$. By such a procedure, regional outputs can be predicted by a means other than allocating national production. The values of the estimated and survey-based tables were compared using a variety of tests, and some encouraging results were obtained.

First, the survey-based table was taken as the true depiction of the 1967 Washington economy. Each of the NAIVE, EMPLQ, and OUTPLQ direct coefficients matrix estimates was compared with the Washington matrix (WASH) and each other. As hypothesized, the NAIVE model proved the least accurate estimate in all but one of the tests. EMPLQ showed substantial improvement in estimation accuracy, with OUTPLQ exhibiting the smallest error of all. Hence, our results indicated that the coefficients displayed a definite increase in accuracy when location quotients were applied to the national table, particularly when output rather than employment was used. For example, from NAIVE to EMPLQ to OUTPLQ, the mean percentage error[5] decreased from 314% to 162% to 135%. However, these raw figures taken by themselves still do not indicate very close estimation. In fact the correlation coefficient ranged, respectively, from .649 to .717 to .733.

The real-life test of the estimated A-matrix is its predictive value, obtaining gross outputs using A and the final demands vector: $X = (I-A)^{-1}F$, where $(I-A)^{-1}$ is the matrix of direct and indirect coefficients. Again significant improvement was evidenced in progressing from NAIVE to EMPLQ to OUTPLQ.

First, in testing $(I-A)^{-1}$, the statistics again showed some variation between estimated and true coefficients. For example, mean absolute difference, reduced by a factor of about two-thirds in moving from NAIVE to OUTPLQ, measured .0056, on the same order as its value for the A-matrix (.0043). The very high correlation coefficients -- .976, .991, and .993 -- are likely due to the extremes of coefficient values, concentrating either at small fractions or greater than one on the matrix's diagonal, allowing a least squares regression to explain much of the variation about the mean.

The ability of the estimated tables to predict actual total output with all final demand given was quite impressive. Mean percentage error (and standard deviation) decreased from 161.95% (239.28%) for NAIVE to only 32.79% (53.59%) for EMPLQ and 16.63% (17.17%) for OUTPLQ. All three estimates overstated total output and a least squares regression line fit very closely, with correlations of .965, .988, and .994 with slopes fractionally greater than one, possibly suggesting some type of bias.

---

[5] Defined in Chapter 3.

Tests of estimated intermediate output AX were also made, resulting as with the coefficient matrices in significant improvement from NAIVE to EMPLQ to OUTPLQ, but with substantial error present. Exports calculated as residuals were also highly correlated with the true values, although highly understated, again a possible indication of some sort of bias. The previously described balancing procedure to insure internal consistency, while not providing drastic changes in the results, does in effect virtually nullify the influence of location quotients in the regionalizing of the national table and so will not be discussed here.

Conclusions and Summary --

In this description of our effort to examine the ability of two types of location quotients to adjust national input-output tables to reflect intraregional interindustry transactions, we have attempted to highlight our efforts at improving the methods used in previous studies.

First, the input-output matrices estimated using location quotients were shown to be a significant improvement over the naive assumption of national interindustry relationships within the region.

Second, location quotients based on output were shown to provide better adjustment factors for the national table than did those in which employment was the measure of importance of an industry.

Third, the application of a large number of previously-used tests of the accuracy of the estimated table allowed some degree of comparison to be made with previous tests. In this way, for example, the effects of a large number of sectors (and hence more homogeneous industry) and the use of output as opposed to employment in computation of location quotients could be examined. It was found in virtually all tests that output data provided better results than employment. And, while in general the results of the present study indicated improved estimation ability over previous studies, contrary conclusions based on some comparisons or other tests may be explained by the increased cumulative error in the summary comparative statistics due simply to the increase in data points (number of sectors, squared).

Finally, in the desire to estimate accurately regional outputs using a short-cut method to regionalize national input-output tables, the most important goal is to be able to accurately estimate regional outputs. Our application of location quotients achieved quite good results using the given Washington final demand vector.

By removing some handicaps suffered by other studies, and achieving improved results, it is tempting to suggest that further improvements may be attained if the problems we could not remove were eventually eliminated. Specifically, definitional differences in sectors and conventions between

Washington and national tables may account for additional error. The transfers of secondary products and imports within the national table in particular could be expected to influence the effect of location quotient adjustments. Further, the consistent overstatement of all three estimates suggests some possible bias. It may be feasible to apply some sort of adjustment factors to account for this bias. While these problems have not yet been solved, the success achieved in the present study suggests that location quotients can indeed be viable means of regionalizing national input-output tables.

When an individual state or region wants to explore the effects of national economic trends, environmental standards or energy development on the economy and environment of its state or region, the preparation of a state or regional I-O table by the location quotient method may be warranted. While the resulting I-O table may not accurately reflect a survey-based regional table, it may be sufficiently close to permit approximate estimates of the direct and indirect effects of exogeneous developments on the regional economy. Thus the location quotient method may provide an additional tool that allows a first order estimate of effects that an elaborate interindustry, interregional model (along with current and projected data bases) could provide if it were available. In other words, location quotients would be, in our view, an acceptable method for regional analysis from the perspective of a regional or state administration, until such time as a satisfactory regional model becomes available.

B. A Review of Alternative Regional and Interregional Models

We have reviewed a variety of existing regional models with the objective of determining which of these models can be incorporated, with or without modification as the Regional Module of SEAS. Due to the complexity of the problem, and the not too infrequent lack of reliable data, most regional models are 'linked' in some fashion to a national model. Examples of these models are the (1) Shift-Share, and (2) Location-Agglomeration Models, (3) Regional and Interregional Input-Output Models, and (4) Regional Macro-Econometric Models. Each of these models is designed to explain and project regional economic activity. At the original time of their formulation, however, these models were not designed for the purpose of analyzing the economic and environmental impact of alternative government energy policies.

Existing Regional Forecasting Models --

The existing regional forecasting models can be classified into four relatively distinct categories and these are: 1) the OBERS "shift-share" type model; 2) the Harris "Location-Agglomeration" type model; 3) the Klein-Glickman type macro-regional econometric models; and 4) the Polenske MRIO model. These categories are not, however, entirely distinct since some of the models utilize a number of different techniques to conduct forecasting analysis. For example, the Harris model and the OBERS projections combine regression analysis with the regional shares, and I-O approaches to determine the effect of alternative transportation systems and water resource development on a regional economy, respectively. MRIO is strictly an Input-Output technique developed to project regional and interregional economic activity, and the Glickman model is strictly an econometric type model using regression analysis to project economic activity in the Philadelphia region.

Energy and the Environment --

None of the models, however, contains explicit formulations to project the environmental impact and the effect on energy demand of alternative energy scenarios. The Glickman model and MRIO are 'pure' regional economic impact models. The OBERS model tries to project the environmental impact of water resource development on land use, and the Harris model introduces environmental quality variables to explain regional variation in output.

Structure of the Models --

All of the models, except MRIO, consider the national economy as the 'driving force behind the regional economies.' The OBERS projections and the Harris model project regional shares of national values - output, employment, income - into the future. Changing shares over time reflect the dynamic character of the model and the shifting location of regional industries. The Glickman regional model of the Philadelphia SMSA was designed to be part of a satellite system with the national economy in the center of the system. MRIO, on the other hand, was designed to estimate each region's level of output, employment, income and expenditures, and then obtain the national totals as the sum of all the regional values.

Geographical Unit of Analysis --

To project economic activity levels OBERS bases its analysis on the BEA Economic Area. The Harris model has also been applied to the 173 Economic Areas, but was initially intended for the county subdivision. The MRIO region is the state, a political unit. The Glickman model is based on the Philadelphia SMSA. Of these modular regions, the MRIO state is probably the most open since SMSA's extend beyond political boundaries. The BEA economic areas are the most self-sufficient in the employment sense since they usually contain labor's place of residence and place of employment in the same area.

C. A Multi-Regional Model of the Regional Economy and Energy Demand (MREED)

In view of the significant drawbacks of the OBERS model currently incorporated in regional forecasting in SEAS, the need for an improved regional forecasting tool in REGION became clear. A survey of existing regional models -- regional and interregional input-output models, regional econometric models, location-agglomeration models -- suggests their limited usefulness for the specific regional information for environmental and energy policy analysis and the computational needs of SEAS.

In the SEAS policy context, it is important to understand the regional economic and environmental effects of alternate scenarios. How do the various regions fare in attracting different primary and manufacturing industries? How will they be affected when different pollution abatement policies are implemented? How do the increased costs of abatement, or changing energy prices affect regional comparative advantage? Which industries and regions are most likely to be sensitive to these changes? What are the effects of all of this on regional income, consumption, local government revenues and taxes? What are the likely consequences on unemployed in the region? What are the likely consequences of energy price increases on consumption of residential, industrial, and transportation energy in various regions? What are the feedback effects of energy price increases on regional growth? What are the likely consequences of large scale energy development in one region (such as ORBES) on the economy and environment of that region and on those of other regions? These are the kinds of questions that a Region Module of SEAS needs to address and the models reviewed above cannot satisfactorily address. Regional models designed for different objectives do not serve SEAS' purposes.

As a direct consequence of the above finding, we have outlined a Model of the Regional Economy and Energy Demand (MREED) whose structure and data output are tailored to the regional information needs of environmental and energy policy analysis. MREED incorporates the key elements of regional economic theory and the more attractive features of existing regional models in a framework of economic-environmental-energy analysis. Large portions of MREED's structure have been previously estimated by the senior author and the relevant data base is available. The proposed incorporation of MREED in the REGION module will follow the general SEAS strategy of improving existing models for incorporation into various modules of SEAS.

Chapter 5 outlines a proposal of continuation of the Johns Hopkins University work on the REGION module directed at developing the core components of MREED. It provides a description of the proposed full structure of MREED. T. R. Lakshmanan has developed and used at the U.S. EPA an earlier regional model for the 100 largest SMSA's.[6]

While the model structure, data base, and research strategy are detailed in Chapter 5 a few salient features of MREED may be highlighted.

1. MREED is a multi-industry, multi-regional model of the U.S. that provides a description of the evolution of various regional economies (BEA economic areas) in the national economic setting. It is formulated as a regional module of SEAS, so that the regional estimates obtained from the model are normalized to the national totals obtained from SEAS. MREED can be used in two ways: First, it will help allocate national (INFORUM) estimates of output by sector to regions. Second, using observations over 10-15 years, the model can be estimated also for each BEA area so that each region's output by sector, income, consumption, government expenditures, labor force, employment population,

---

[6] T. R. Lakshmanan and Fu-Chen Lo. "A Regional Economic Model for the Assessment of Effects of Air Pollution Abatement," *Environment and Planning*, Vol. 4, 1972, pp. 73-97.

and energy demand by major end use can be estimated. Thus, the model could be used for assessing economic and environmental and energy effects in the SEAS framework either for one region such as ORBES Region or for allocating national totals to regions. Since MREED will produce economic information at the BEA area level, allocation of some sort to the county level is necessary to permit the flexibility in regional area data in SEAS studies.

2. MREED is a 26-equation econometric model organized at the level of 173 BEA areas -- each of which is a labor market area. It is organized into four blocks -- output, income and expenditures, labor market and energy. The equations in those blocks incorporate principles of regional location theory and interregional interaction theory in a framework where environmental quality and energy prices affect comparative regional advantage. Thus, output in manufacturing and primary sectors is determined in national markets (viewed multi-regionally) under the influence of regional environmental quality and different regional energy prices. Energy demand is in turn related to income, industrial structure, and price. While the resulting structure of the model is elaborate, it must be noted that the greater portion of the model has been already estimated (in the form presented here or quite close to it) by the principal investigator and others (Lakshmanan and Lo 1972, Harris 1973, Glickman 1974). The extension of this previous work embodied in MREED consists of adding and incorporating environmental and energy variables as factors affecting region location of industries and households and in identifying energy consequences of regional economic growth. Thus the developmental work in MREED is far more limited than the model structure described in the next section would suggest.

3. The strategy of development of MREED should follow the evolutionary approach and be organized in phases. The first phase can be visualized as the development of only two blocks -- manufacturing output, and energy -- at a sector detail for which BEA area data are readily available. This work will permit the assessment of the role of regional environmental quality and energy input costs in industrial location and the level of regional energy demand -- industrial, residential, and transportation. The information developed here will greatly augment the utility of the regional module of SEAS. In the next phase, the remaining blocks can also be estimated at the BEA region level.

D. A Model to Assess Regional Economic Effects of Final Demand or Energy Price Changes

Despite the limitations of input-output analysis and the resource constraints involved in collecting national and regional data, the model developed by Miyazawa and generalized by Yamada and Ihara and discussed in Chapter 6 will prove to be an invaluable instrument to analyze national, regional, and interregional economic repercussions of alternative energy and public expenditure policies.

This model can be used to analyze the effects on regional output, income, private expenditures, and interregional trade flows of an exogenous change in public expenditures in either of the two regions.

Furthermore, we can analyze the effect of an exogenous change in a region's vector of prices. The impact will first be felt in that region's level of income, and will then influence output, private expenditures and interregional trade flows through the income multiplier process.

This model can also be used in an entirely different way. The two regions can be viewed as two distinct industrial sectors -- industrial goods and agricultural goods producing sectors, energy-intensive and non-energy intensive sectors, etc. As with the analysis in the two region case, the effect on output, income, prices in a given sector, and intersectoral flows can be analyzed. By categorizing the industries in a national I-O table into energy producing and non-energy producing industries, for example, we can examine the degree of dependence (independence) of non-energy producing industries on the energy sector. This can be done by examining how non-energy output and intersectoral trade flows respond to an expansion in the energy sector. Assuming, however, fixed production and trade coefficients, one must realize the dependence of one sector on the other will be overstated.

Similarly, 'cost-push' inflation caused by increased energy prices can be analyzed. Again by dividing the national I-O table of prices and wages into those pertaining to energy and non-energy producing sectors the effect of energy price increases on other prices can be determined. However, we must again realize that by ignoring the price sensitivity of inputs, the extent of 'cost-push' inflation caused by energy price increases will be overstated.

CHAPTER 2

AN ASSESSMENT OF THE DATA BASE OF REGION

This chapter presents the results of an assessment of the data base used in the Regionalization (REGION) Module of the Strategic Environmental Assessment System (SEAS) Model. It opens with a brief summary of the REGION Model and its assumptions. We proceed then to a comparison of the state output data developed by SEAS for 1972 with the state output data for the same year developed by the U.S. Bureau of Census for 34 INFORUM sectors in 15 major states. We interpret the differences in regional output between these two sources in terms of sector productivity variations at the state level. From this empirical exercise, we draw some inferences for future directions in strengthening the regional data base in REGION.

I. THE REGION MODEL

The purpose of REGION is to allocate national economic output, abatement costs and pollutant emissions developed by other SEAS modules (INFORUM, ABATE, RESGEN) to various regional units. REGION does not incorporate an analytical description of the structure of production and consumption in the regions or the economic interactions among regions in the country. Instead, it provides for a straightforward proportional allocation of a base year national sector output to regions based on the latter's proportional share of national employment in that sector.

$$X_{ijt} = X_{it} \frac{L_{ijt}}{L_{it}} \qquad (1)$$

where X and L represent output and employment in industrial sector i, j is the regional unit and t is the time period.

The use of the more readily available regional data on employment in Equation (1) implies an assumption of equal output per worker in a sector in all regions. (In some instances, plant capacity data are substituted for employment in Equation (1).)

For a base year, the REGION model develops 18 sets of regional fractions termed SHARE (= L_{ijt}/L_{it} in Equation (1)) to be applied to the national estimates to determine regional output and pollutant emissions. Nine of these SHAREs are intended to regionalize the output and abatement costs to any of the following regional levels:

1.	EPA Federal Regions	(10)
2.	States	(51)
3.	Standard Metropolitan Statistical Areas (SMSA)	(254)
4.	Air Quality Control Regions (AQCR)	(247)
5.	Major River Basins	(19)
6.	Minor River Basins	(221)
7.	Aggregated Subareas (ASA)	(101)
8.	Bureau of Economic Analysis Regions (BEA)	(173)
9.	Non SMSA BEA Regions	

For each of the above, E-SHARE is computed as follows:

$\text{E-SHARE}_{ij}(t) =$ the proportion of output and abatement activity associated with sector i in regional unit j in year t.

The other 9 SHAREs are for regionalizing pollutant emissions to any of the above 9 regional units and are computed as follows:

$\text{P-SHARE}_{ij}(t) =$ the proportion of pollutant emissions derived from sector i in region j in year 6.

The definition of P-SHARE in this manner implies that the national shares of all pollutant types associated with a sector in any one region are the same. (This is modified in selected sectors, however.)

The E-SHARE and P-SHARE files are developed for a base year (1971) using employment data (as indicated earlier, for a few sectors plant capacity data are used) at the county level and aggregated to the desired regional level. Thus

$$\text{SHARE}_{ij}(1971) = \sum_{c=1}^{m} L_{ic}(1971) \Big/ \sum_{c=1}^{n} L_{ic}(1971) \qquad (2)$$

where c refers to a county and there are m counties that make up a regional unit j and n counties make up the nation.

Economic Information System (EIS) is the source of REGION's base year data on $L_{ic(t)}$. EIS is a private company deriving its data for 115,000 industrial establishments by name and address, state and county, and by 4 digit SIC. Since these data are drawn from public sources, there are no disclosure problems. The data base, also known as the "Gould Tape" (after the President of EIS) is considered a reasonably complete survey of all plants in the continental U.S. with more than 20 employees. EIS claims completeness of data for plants with more than 250 employees, 98% accuracy in the range of 50-250 employees and 85% to 98% for plants with 20-50 employees. It does not cover plants with less than 20 employees.

The SHARE values are obtained for future years by utilizing the 1972 Office of Business Economics (OBERS), Projections of Economic Activity in the U.S. These projections of earnings put out by the U.S. Department of Commerce at the 2 digit SIC are used to compute an OBERS growth factor ($OG_{i',j}$). The latter is defined as follows:

$$OG_{i,j}(t) = OBERS_{i,j}(t) \Big/ OBERS_{i,j}(1971) \qquad (3)$$

where $OG_{i,j}$ is the OBERS growth factor associated with the 2 digit SIC sector i' in regional unit j. $OBERS_{i,j}$ is defined as the industry earnings associated with i' in region j.

For each INFORUM sector included in the OBERS sector i', the projection year (t) SHARE values are computed:

$$SHARE_{ij}(t) = SHARE_{ij}(1971) \times OG_{i,j(t)} \qquad (4)$$

The SHARE values in Equation (4) are normalized over all regions in sector i to sum to 1. Thus the growth patterns in each of the 3 or 4 digit INFORUM sectors are constrained to that of the 2 digit OBERS sector they are contained in. OBERS makes projections for 37 economic sectors compared to INFORUM's 185.

The allocation procedures used in REGION have a number of drawbacks. The use of employment as a surrogate for output in REGION is acknowledged as introducing errors whenever output per worker levels vary significantly among regions, but was judged in the design phase of SEAS as not producing unreasonably great inaccuracies. Output per worker within a sector does vary considerably in response to a number of factors, such as process or product mix. In an industry with high capital investment and relative "stability" in the markets for the product, older plants are likely to continue producing in the face of competition from newer plants using different processes and capital/labor ratios. While each plant produces essentially the same type of product, their production processes differ. If that sector has experienced regional location shifts (such as textile plants moving from the Northeast to the South) differences in plant age, and processes may be highly correlated with specific regions. Another problem is that the type of products classified under a single INFORUM sector (and indeed under any classification scheme with a reasonably limited number of sectors) may differ widely. "Apparel" produced in New York may involve expensive fashionable clothing with relatively high value added compared to the work clothing produced in Wisconsin. Apparel in Texas may be relatively labor intensive, taking advantage of the availability of less expensive labor. The point made here is that output per worker is sufficiently varied for a number of sectors that the number of employees is a poor indicator of levels of output.

Second, the OBERS method of projecting regional growth to be built around a "shift share" model which extrapolates past locational patterns.

$$LS_{ij} = E_{ijt} - E_{ijo}\left(E_{it}/E_{io}\right) \qquad (5)$$

where LS_{ij} = locational shift of industry i in region j between time periods o and t.

E_{ijt} and E_{ijo} are levels of economic activity in sector i, region j at time periods t and o.

E_{it} and E_{io} are national levels of economic activity in sector i at time periods t and o.

OBERS trends LS_{ij} into the future with the result that future locational shifts mirror only past ones and do not take into account the effects of dynamic factors on changes in regional comparative advantage.

Regional pollution emissions are obtained in SEAS on the assumption that a region has the same share of national emissions as it has of the national output. This assumption ignores differences in process mix, industry mix and environmental standards, plant age, etc., among regions, but still serves as a useful first approximation.

The objective of the study reported in this chapter is a limited one. It is not to improve the workings of REGION model structure, but to suggest improvements in the economic data base, so that regional shares (and in turn pollutant estimates) can be made more reliable. For this purpose, the comparison of a base year SEAS regional data base with an independent regional data base is made in the next section.

II. COMPARISON OF REGIONAL DATA BASES

Regional estimates of economic output for 34 INFORUM sectors were compared for a sample of 15 large states with data from the *1972 Census of Manufactures* published by the Bureau of the Census. The *Census of Manufactures* involves a complete and detailed survey of all manufacturing establishments with 20 or more employees, and less detailed data on plants with less than 20 employees. It provides data down to the 5-digit Standard Industrial Classification (SIC) code for individual states, SMSA's and other regions, including counties.

However, the CM data is considered confidential and so disclosure limitations preclude the publication of data which would reveal specific information about an identifiable plant. Obviously, at the level of the county, disclosure problems become more significant. Because the most complete sector detail is published at the state level, the present study chose this level as the basis of comparison with SEAS. The CM data gathering effort is extensive and the publications include data on a wide variety of industry related subjects. The comprehensiveness of the effort also means that the data is made available several years after the period of interest (1972 data was published in book form in 1976). Of the 121 INFORUM sectors (Numbers 20-150) for which CM data are available, 34 sectors were chosen for this comparison.

A directory of 1972 SIC sectors comparable to given INFORUM sectors was obtained from the U.S. Environmental Protection Agency. The Value of Shipments for each SIC sector in a given state included in an INFORUM sector was compared with the value of output provided in a SEAS Base Case Run.

It should be noted that the use of the Census of Manufactures data as "ground truth" may be open to some question. As in any large scale data gathering effort, sampling and aggregation errors are probably inevitable. The Census of Manufactures does, however, have the advantage of a long period of data gathering and preparation and of being the most recent effort of a data series which has been produced for a number of years. In contrast, the EIS Gould Tape was used much sooner after the period under study, and the 1971 data was one of the early ones undertaken by EIS. Further, the Gould Tape aggregates from the county level, a relatively small geographical area, which increases problems of randomness in the data and the possibilities of aggregation errors particularly due to incomplete coverage of smaller firms. Such factors certainly do not ensure the greater accuracy of the Census of Manufactures, but are in its favor as a benchmark. Thus while the text will refer to the "accuracy" of the SEAS data, such references are premised on the validity of the CM figures.

What follows is a description of the procedures used to determine how closely SEAS estimates correspond with data from the Census of Manufactures; a description of the degree of correspondence by sector and state; an analysis of output per worker differences as an explanatory factor for differences and the development of a modifier based on the output per worker variations which could be determined; and an appendix listing the specific information used in each sector.

Several discrepancies between SEAS and CM conventions required adjustment before a comparative analysis could be made. SEAS INFORUM projections are classified according to a "purified" sectoring scheme -- a sector's outputs consist only of that product. However, CM data on a state level are only available classified at an establishment level -- all outputs of an establishment are assigned to a sector if its production of that sector's products is greater than that of any other sector's products. National CM data are available on a "purified" basis,[1] however, so a "purification correction factor" -- the ratio of Purified National CM Value of Shipments to Unpurified National CM Value of Shipments was determined for each INFORUM sector (Table 2-1). This ratio was then used to modify the available state CM data to correct for the lack of purified figures. (This adjustment implicitly assumes the national ratio of "purified" to "unpurified" production in an industry holds in each state. However, it was seen as the only means of making the adjustment with the data available.)

---

[1] Census of Manufactures, Industry Series, Tables 5-B. Purified Values of Shipments are obtained from addition of an industry's primary products and miscellaneous receipts.

It would have been preferable if the SEAS base year (1971) coincided with the data published for every fifth year (1967, 1972, 1977) by the CM. Thus it was necessary to compare a projection year estimate (1972) by SEAS with the CM data. First, as SEAS output is reported in 1971 dollars, adjustments for inflation were made by multiplication of SEAS projections by appropriate price indices (Table 2-2). Secondly, since the purpose of this research effort is to validate base year data rather than to check the accuracy of projections, an attempt was made to reduce errors due to "inaccurate" projection. Within SEAS, the 1971 base year totals agree with known levels and are used as controls for the regionalization (i.e., the sum of all the regional outputs in a given sector must equal the known national total). The regional output data of SEAS were summed for each sector in 1972 and were compared with the totals published in the CM. The SEAS state-level estimates were then modified by national sector scaling factors which reflect the differences between the SEAS-estimated sector growth and the CM sector growth in the nation for each sector (Table 2-3). Of course, REGION uses regional differential growth rates from OBERS and so a national correction factor would not completely correct for the projection inaccuracies in each region. However, since only a projection of one year is involved, such considerations are believed to be very minor.

The 34 sheets presented at the end of this chapter display the comparisons between CM state output data and SEAS 1972 data modified for national sector total differences in 1972 between CM and SEAS. In addition, the tables indicate the percentage of national output accounted for by each state in the sample, and the national and state output per worker for each of the 34 sectors.

Table 2-4 summarizes, for 5 major states, the SEAS-CM comparisons for 34 sectors displayed in these Appendix tables. Substantial variation between the estimated (SEAS) and true (CM) values can be seen. Mean statewide SEAS/CM ratios range from .95 for California to 1.31 for Texas. Variation about the means is large as well; standard deviations for the five states are .20 (Illinois), .23 (California), .30 (Ohio), .37 (New York, and .63 (Texas) over the 34 ratios.

III. OUTPUT PER WORKER MODIFIER

Output per worker (O/W) levels were examined to determine if differences in such levels could explain all or part of the observed SEAS/CM differences.

The O/W was calculated from the CM data for each sample state for each of the 34 INFORUM sectors. These values were then compared with national

TABLE 2-1. PURIFICATION OF CENSUS OF MANUFACTURES VALUES OF SHIPMENTS

INFORUM Sector	(1) Census of Manufactures Value of Shipments ($ millions)	(2) Purified Value of Shipments ($ millions)	(2)/(1) Purification Correction Factor
23	31,447.9	38,454.7	1.22
24	16,311.5	15,998.2	.98
25	11,478.6	11,295.2	.98
26	12,162.2	13,011.4	1.07
27	7,895.5	7,943.8	1.01
31	6,925.8	7,046.1	1.02
39	22,709.7	24,500.4	1.08
40	5,099.5	5,887.8	1.15
45	7,409.6	7,385.0	1.00
51	8,104.7	8,148.1	1.01
52	8,262.8	8,200.0	.99
53	3,510.6	3,584.7	1.02
54	3,794.5	4,140.4	1.09
59	2,778.3	2,909.8	1.05
66	8,018.5	7,830.4	.98
67	9,777.6	9,998.4	1.02
69	26,761.8	28,163.4	1.05
76	3,478.7	3,486.1	1.00
81	9,537.6	9,398.0	.99
84	8,432.4	12,734.6[a]	1.51
89	1,244.6	1,197.7	.96
94	2,136.8	2,043.3	.96
95	14,179.6	14,525.4	1.02
97	8,318.3	8,351.1	1.00
98	5,094.3	5,111.2	1.00
101	4,288.3	N.A.[b]	(1.00)
103	5,590.4	5,501.7	.98
107	696.3	709.7	1.02
113	1,813.1	1,842.4	1.02
118	1,540.8	1,408.6	.91
119	3,585.4	3,605.2	1.01
123	6,938.6	6,973.6[c]	1.01
127	13,665.0	13,353.6	.98
128	8,836.6	9,038.0	1.02

[a] Extensive duplication in shipments.

[b] NA=Not available.

[c] Small undisclosed miscellaneous receipts entry not included.

TABLE 2-2. 1972 PRICE INDICES FOR INFORUM SECTORS (1971 = 1.00)

INFORUM Sector	Price Index
23	1.12
24	1.03
25	1.05
26	1.04
27	1.03
31	1.02
39	1.02
40	1.05
45	1.02
51	1.03
52	1.02
53	1.05
54	1.02
59	1.01
66	1.01
67	1.00
69	1.02
76	1.07
81	1.05
84	1.04
89	1.02
94	1.03
95	1.04
97	1.05
98	1.03
101	1.05
103	1.04
107	1.02
113	1.02
118	1.01
119	.99
123	1.00
127	1.00
128	1.00

SOURCES: Wholesale Prices and Price Indexes (1971, 1972), Bureau of Labor Statistics.
Handbook of Labor Statistics (1974), Bureau of Labor Statistics.

TABLE 2-3. SELECTED NATIONAL SECTOR OUTPUTS FOR 1972

SEAS SECTOR #	PURIFIED C.M. ($1972)	SEAS ($1972)	$\frac{CM}{SEAS}$ = Scaling Factor
23	38,454.7	25,557.9	1.50
24	15,998.2	16,095.6	.99
25	11,295.2	12,610.5	.90
26	13,011.4	9,974.4	1.30
27	7,943.8	7,861.0	1.01
31	7,046.1	6,362.9	1.11
39	24,500.4	23,558.1	1.04
40	5,887.8	4,712.2	1.25
45	7,385.0	7,565.4	.98
51	8,148.1	7,820.3	1.04
52	8,200.0	7,982.4	1.03
53	3,584.7	4,563.6	.79
54	4,140.4	3,641.0	1.14
59	2,909.8	1,900.4	1.53
66	7,830.4	7,642.3	1.02
67	9,998.4	9,283.6	1.08
69	28,163.4	27,975.3	1.01
76	3,486.1	4,015.8	.87
81	9,398.0	9,036.6	1.04
84	12,734.6[a]	6,342.5	2.01
89	1,197.7	978.1	1.22
94	2,043.3	1,959.1	1.04
95	14,525.4	14,105.0	1.03
97	8,351.1	7,741.3	1.08
98	5,111.2	5,206.2	.98
101	4,288.3[b]	5,052.0	.85
103	5,501.7	4,460.8	1.23
107	709.7	772.3	.92
113	1,842.4	2,001.3	.92
118	1,408.6	1,308.3	1.08
119	3,605.2	3,613.6	1.00
123	6,973.6	7,611.9	.92
127	13,353.6	15,372.5	.87
128	9,038.0	9,050.6	1.00

---

[a] See notes in Table 2-1.

[b] Value of shipments not purified due to unavailability of data

TABLE 2-4. SEAS/CM FRACTIONS FOR SELECTED STATES[a]

SECTOR	NY	OHIO	ILL	CAL	TEX	SECTOR	NY	OHIO	ILL	CAL	TEX
23	2.03	1.77	1.34	.98	1.26	76	1.00	D	.77	.73	.66[b]
24	.57	1.12	.60	.74	2.49	81	D	1.16	1.23	.85	.98
25	.74	.68	.98	.86	1.23	84	.83	1.67	1.10	D	D
26	.97	.90	1.23	.78	1.38	89	.45	1.27	D	D	D
27	.94	.93	.87	.92	1.45	94	.96	.99	1.20	D	D
31	.74	1.18	1.13	.66	D	95	1.20	.89	1.19	.95	1.14
39	.43	D	1.31	D	D	97	D	.88	1.08	1.51	3.02[b]
40	1.13	.61	1.03	1.20	1.51	98	1.04	.84	1.00	.90	2.41[b]
45	1.42	.79	1.33	.80	1.02	101	D	D	.83	D	D
51	1.17	.98	1.15	.75	.85	103	1.09[b]	1.02	1.16	D	1.77[b]
52	.98	1.07	1.11	.90	1.08	107	1.08	.70	1.03	D	.72[b]
53	.62	1.72	1.09	.61	D	113	.90	D	1.35	D	1.12[b]
54	.79	1.06	.99	1.16	2.11[b]	118	1.03	1.17	.96	D	.69
59	1.21	D	D	D	.61	119	1.56	.95	1.31	1.00	.90
66	.62	1.06	1.32	1.28	1.50[b]	123	D	1.38	1.03	.75	.55
67	1.71	1.29	.77	1.39	1.18	127	1.06	1.13	.98	.99	D
69	D	.61	1.41	.93	1.12	128	.50	.93	1.29	1.16	D

[a] SEAS is modified by the price index and **scaling** factor; CM is modified by the purification correction factor. No modification has been made for O/W differences.

[b] State accounts **for less than 1% of total sector output.**

D = Disclosure limitation

O/W values for each sector.[2] If the differences between SEAS and CM are to be explained by differences in the O/W among the states, an O/W for a state which is higher than the national average for that sector should lead SEAS to proportionally underestimate production in that state, and a lower than national average O/W for a state should lead SEAS to proportionally overestimate production. The ratio of state to national O/W was used to adjust the SEAS/CM fraction to test this hypothesis; the adjusted SEAS estimates would then approach those of CM (and SEAS/CM would approach 1.00). For the most part, the trends in the state O/W and the corresponding differences in the SEAS/CM fraction tended to follow the expected path. However, there were frequent cases where O/W failed to explain the full difference in the SEAS/CM.

Although the differences in output per worker among the sample states does not always correct for the differences in the SEAS/CM figures, it still provides a very useful explanation for much of the observed differences. Besides OBERS growth, a second modifier within SEAS based on the variation in output per worker would serve to bring most of the SEAS/CM figures more closely in line. Table 2-5 demonstrates the effect of this second modifier for sample states in 3 sectors. The sectors were chosen because of the relatively wide variation on O/W levels within each sector. States with disclosure problems or which accounted for 1% or less of sector output are not shown. The ratio of state O/W and national O/W is multiplied by the SEAS/CM fraction. If the state has less than a national average output per worker level, this modifier will be less than one and so the SEAS/CM fraction will be deflated. Although some SEAS/CM fractions actually move further from 1 (the standard of perfect agreement), most of the fractions moved closer to 1 when multiplied by the O/W modifier, and a number showed significant improvement. For example, the difference between upper and lower values for sector 23 narrowed from 1.4 to 0.9, with standard deviation of .44 around a mean of 1.24 dropping to .32 about a mean of 1.22. In sector 26, the mean (and standard deviation) improved from .955 (.27) to 9.24 (.18). For sector 65, the change was from 1.04 (.31) to 1.01 (.21).

If state level output data were available without disclosure problems it would be relatively simple to incorporate the differences in output per worker into the regional allocation factors. Given the disclosure limitations it is not possible to correct for O/W differences in all states in all sectors. One issue then is whether to use available data on O/W differences or not to use it because it is not complete. One possibility is to perform a two-part allocation. In the first part states would receive their share of a sector's output based on employment modified by state level output per worker. The sum of output of all states for which the O/W was available could then be subtracted from the national control total, and the remaining states would receive allocations from the remaining national output. On the basis of the work performed for this project, it is estimated that O/W levels could be developed for between one half and two thirds of all states in each of the 121 SEAS industrial sectors.

---

[2] O/W fractions based on unpurified values of shipments were used, but both state and national output are purified by multiplication by the same purification correction factor for a given sector, so for comparative purposes, no conversion was required.

Of course, the calculation of state level output per worker levels would only be useful if the differences among the states were relatively stable over the projection period. It would seem that for a period of up to a decade the regional differentials in O/W may not vary to a significant extent. If the output per worker differences are due to differences in product and process mix, the forces which affect this mix would tend to act over relatively long time periods. This assumption needs to be verified by an analysis of past Census of Manufactures data before this approach can be seriously advanced.

IV. IMPLICATIONS

It appears that the SEAS output data at the state level obtained using the Gould tape display considerable differences with CM data for quite a few sectors and states. Apparel (#39) and electronic components (#128) are examples of INFORUM sectors in which the output estimates for a few major producing states are widely divergent from the CM data.

Some of these differences are attributable to the significant departures of output per worker in those states from the national average. Corrections for such differences bring the SEAS and CM estimates closer. The remaining differences, assuming the veracity of CM data, may be attributable to the fact that SEAS aggregates state data from county level data. The "noise" at the fine geographical detail of a county is another source of potential differences.

There is an obvious trade-off between flexibility and accuracy in the choice of the county as the basic building block. The use of the county as the basic unit permits flexible assembly of regional data in any one of 9 different regional units. However, data bases and modelling of locational and interregional relationships are more tractable at a larger regional level such as a state or BEA economic area.

Consequently, it may be necessary to explore an approach complementary to the current regionalization work in SEAS. What we would like to consider is an assembly of a base year economic output data base at the state level for the INFORUM sector. Such a calibration of base year regional data base will utilize the best available output data. The base year could be a census year such as 1972 and CM data can be supplemented with data for other sectors. Where output data are not available because of disclosure problems, other sources such as EIS can be used in a supplementary fashion. Such a careful inventory of economic data base will increase the credibility and usefulness of SEAS regional information.

TABLE 2-5. EFFECT OF O/W MODIFIER ON SEAS/CM FRACTIONS FOR SELECTED STATES AND SELECTED SECTORS[a]

SECTOR	STATE	$\frac{\text{SEAS}}{\text{CM}}$	$\frac{\text{SEAS}}{\text{CM}} \times \frac{\text{STATE O/W}}{\text{NATL. O/W}}$
#23	NY	2.03	1.57
	NJ	.89	.85
	PA	1.47	1.33
	MICH	.59	.79
	OH	1.77	1.58
	ILL	1.34	1.44
	WIS	1.41	1.47
	NC	1.24	.67
	TEX	1.26	1.17
	COL	.66	1.34
	CAL	.98	1.18
#26	NY	.97	1.05
	PA	.86	.84
	MICH	1.31	1.02[b]
	OH	.90	1.00[b]
	ILL	1.23	1.02[b]
	WIS	.50	.62
	NC	.78	1.00[b]
	ALA	.84	1.12
	TEX	1.38	1.31
	CAL	.78	.96[b]
#66	NY	.62	.73
	NJ	.96	1.04
	PA	.75	1.05[b]
	MICH	1.29	1.36
	ILL	1.32	.98[b]
	CAL	1.28	.91

---

[a] These 3 sectors were chosen as illustrations because O/W among states varies considerably in each. The number of states shown vary from sector to sector because of disclosure problems. Also states with 1% or less of sector output are not shown.

[b] Significant Improvement in SEAS/CM.

V. APPENDIX -- DETAILED TABLES FOR EACH OF THE 34 SECTORS

The following tables list for each of the 34 INFORUM sectors examined, individual state figures for 1972 output according to the Census of Manufactures, those projected by SEAS, and their ratios. Output per worker and percentage of national output are also listed, as are national summary figures and adjustment factors (described in the text) for each sector.

The following terms are used in the tables. The national purification correction factor was used to adjust the Census of Manufactures data to conform to INFORUM's sector definitions. The scaling factor corrects for SEAS' projection from 1971 to 1972. The 1972 price index is based on 1971 = 1.00. The output per worker calculations were based on unpurified sectors. The purified CM column was calculated by multiplying state Census of Manufactures data by the national purification correction factor. The modified SEAS column was obtained by multiplying the state SEAS figure by the price index (to adjust for SEAS' 1971 dollars) and the scaling factor which accounts for projection error. The appearance of a D in the tables indicates the data was not listed in the Census due to disclosure limitations while NL indicates no listing.

TABLE 2-6. SECTOR #23 MEAT PRODUCTS
Equivalent S.I.C. Sector(s): 201
Census of Manuf. (CM) Nat'l Purified Value of Shipments ($\$x10^6$): 38,454.7
National Sector Output per Worker ($\$x10^3$): 102.2
Nat'l Purif. Corr. Factor: 1.22 Scaling Factor: 1.50 1972 Price Index: 1.12

STATE	SEAS(modified) / CM(purified)	Purified CM Amt. ($\$x10^6$)	Modified SEAS Amt. ($\$x10^6$)	% of Nat'l CM Output	State O/W
Mass.	1.55	405	626.3	1	87.4
N.Y.	2.03	947	1,918.6	2	79.2
N.J.	.89	622	555.4	2	98.1
Pa.	1.47	1,514	2,219.3	4	92.6
Mich.	.59	1,170	696.0	3	137.0
Ohio	1.77	1,299	2,305.0	4	91.0
Ill.	1.34	2,206	2,950.1	6	109.6
Wis.	1.41	1,248	1,765.7	3	106.6
N.C.	1.24	736	914.4	2	54.9
Fla.	1.44	479	691.0	1	78.6
Ala.	1.66	575	952.0	1	50.1
Tex.	1.26	1,974	2,489.8	5	94.6
Ariz.	.61	329	200.4	1	207.7
Col.	.66	1,549	1,023.1	4	208.2
Calif.	.98	2,339	2,299.9	6	122.9

TABLE 2-7. SECTOR #24 DAIRY PRODUCTS
Equivalent S.I.C. Sector(s): 202
Census of Manuf. (CM) Nat'l Purified Value of Shipments ($\$x10^6$): 15,998.2
National Sector Output per Worker ($\$x10^3$): 80.5
Nat'l Purif. Corr. Factor: .98 Scaling Factor: .99 1972 Price Index: 1.03

STATE	SEAS(modified) / CM(purified)	Purified CM Amt. ($\$x10^6$)	Modified SEAS Amt. ($\$x10^6$)	% of Nat'l CM Output	State O/W
Mass.	1.53	325	497.3	2	87.4
N.Y.	.57	1,157	655.1	7	93.0
N.J.	.83	267	222.5	2	90.7
Pa.	1.63	943	1,533.7	6	66.3
Mich.	.76	570	434.6	4	93.9
Ohio	1.12	857	962.1	5	59.5
Ill.	.60	907	543.0	6	95.5
Wis.	.42	2,059	865.3	13	137.3
N.C.	1.21	224	270.4	1	71.6
Fla.	1.14	347	397.1	2	73.7
Ala.	2.05	168	344.6	1	48.9
Tex.	2.49	455	1,130.9	3	62.7
Ariz.	.84	90	75.8	1	92.0
Col.	1.23	114	140.0	1	72.5
Calif.	.74	1,290	959.4	8	105.3

TABLE 2-8. SECTOR #25 CANNED AND FROZEN FOODS

Equivalent S.I.C. Sector(s): 203

Census of Manuf. (CM) Nat'l Purified Value of Shipments ($\$x10^6$): 11,295.2

National Sector Output per Worker ($\$x10^3$): 49.2

Nat'l Purif. Corr. Factor: .98 Scaling Factor: .90 1972 Price Index: 1.05

STATE	SEAS(modified) / CM(purified)	Purified CM Amt. ($\$x10^6$)	Modified SEAS Amt. ($\$x10^6$)	% of Nat'l CM Output	State O/W
Mass.	2.41	71	171.1	1	51.4
N.Y.	.74	779	579.3	7	61.6
N.J.	1.01	527	530.4	5	53.8
Pa.	.94	700	659.8	6	52.1
Mich.	.96	413	397.3	4	50.7
Ohio	.68	517	353.7	5	71.3
Ill.	.98	539	530.7	5	52.9
Wis.	.88	427	374.7	4	41.5
N.C.	1.28	85	109.2	1	51.9
Fla.	.85	917	777.7	8	76.7
Ala.	1.07	60	64.2	1	43.6
Tex.	1.23	340	418.9	3	50.3
Ariz.	--	D			
Col.	2.24	26	58.3	<1	27.0
Calif.	.86	2,504	2,156.5	22	51.1

TABLE 2-9. SECTOR #26 GRAIN MILL PRODUCTS

Equivalent S.I.C. Sector(s): 204

Census of Manuf. (CM) Nat'l Purified Value of Shipments ($\$x10^6$): 13,011.4

National Sector Output per Worker ($\$x10^3$): 109.3

Nat'l Purif. Corr. Factor: 1.07 Scaling Factor: 1.30 1972 Price Index: 1.04

STATE	SEAS(modified) / CM(purified)	Purified CM Amt. ($\$x10^6$)	Modified SEAS Amt. ($\$x10^6$)	% of Nat'l CM Output	State O/W
Mass.	2.20	48	104.1	<1	75.0
N.Y.	.97	584	565.2	5	118.7
N.J.	1.48	83	122.8	1	78.0
Pa.	.86	471	407.0	4	107.3
Mich.	1.31	616	804.8	5	84.7
Ohio	.90	547	490.6	4	121.7
Ill.	1.23	1,464	1,796.8	11	90.6
Wis.	.50	289	143.9	2	135.0
N.C.	.78	330	256.3	3	140.0
Fla.	1.29	107	138.4	1	83.3
Ala.	.84	281	235.7	2	146.1
Tex.	1.38	844	1,163.8	6	103.8
Ariz.	1.36	35	47.6	<1	110.0
Col.	5.68	111	630.7	1	80.0
Calif.	.78	1,056	824.3	8	135.2

TABLE 2-10. SECTOR #27 BAKERY PRODUCTS

Equivalent S.I.C. Sector(s): 205

Census of Manuf. (CM) Nat'l Purified Value of Shipments ($\$x10^6$): 7,943.8

National Sector Output per Worker ($\$x10^3$): 33.5

Nat'l Purif. Corr. Factor: 1.01 Scaling Factor: 1.01 1972 Price Index: 1.03

STATE	SEAS(modified) / CM(purified)	Purified CM Amt. ($\$x10^6$)	Modified SEAS Amt. ($\$x10^6$)	% of Nat'l CM Output	State O/W
Mass.	.88	232	204.8	3	33.8
N.Y.	.94	559	525.4	7	30.7
N.J.	.65	485	313.1	6	40.7
Pa.	1.14	652	740.2	8	33.3
Mich.	.98	307	299.7	4	34.9
Ohio	.93	521	486.4	6	34.9
Ill.	.87	715	621.7	9	38.9
Wis.	1.48	129	191.5	2	27.8
N.C.	1.11	270	300.9	3	27.2
Fla.	.91	186	169.2	2	29.7
Ala.	.78	97	75.6	1	34.3
Tex.	1.45	322	467.0	4	30.7
Ariz.	.84	39	32.8	<1	32.5
Col.	1.07	117	125.6	1	36.2
Calif.	.92	750	688.0	9	35.7

TABLE 2-11. SECTOR #31 SOFT DRINKS AND FLAVORINGS

Equivalent S.I.C. Sector(s): 2086,2087

Census of Manuf. (CM) Nat'l Purified Value of Shipments ($\$x10^6$): 7,046.1

National Sector Output per Worker ($\$x10^3$): 52.75

Nat'l Purif. Corr. Factor: 1.02 Scaling Factor: 1.11 1972 Price Index: 1.02

STATE	SEAS(modified) / CM(purified)	Purified CM Amt. ($\$x10^6$)	Modified SEAS Amt. ($\$x10^6$)	% of Nat'l CM Output	State O/W
Mass.	---	D			
N.Y.	.74	626	463.1	9	73.1
N.J.	.50	346	173.1	5	84.7
Pa.	1.21	300	362.4	4	49.0
Mich.	.72	256	184.9	4	59.8
Ohio	1.18	363	427.4	5	48.8
Ill.	1.13	277	314.3	4	34.0
Wis.	1.38	100	137.6	1	40.8
N.C.	---	D			
Fla.	1.00	240	240.5	3	53.4
Ala.	---	D			
Tex.	---	D			
Ariz.	1.05	50	52.5	1	40.8
Col.	---	D			
Calif.	.66	609	403.3	9	74.6

TABLE 2-12. SECTOR #39 APPAREL

Equivalent S.I.C. Sector(s): 231 through 238

Census of Manuf. (CM) Nat'l Purified Value of Shipments ($\$x10^6$): 24,500.4

National Sector Output per Worker ($\$x10^3$): 19.2

Nat'l Purif. Corr. Factor: 1.08 Scaling Factor: 1.04 1972 Price Index: 1.02

STATE	SEAS(modified) / CM(purified)	Purified CM Amt. ($\$x10^6$)	Modified SEAS Amt. ($\$x10^6$)	% of Nat'l CM Output	State O/W
Mass.	---	D	---		
N.Y.	.43	7,454	3,236.5	30	31.8
N.J.	1.04	1,152	1,194.4	5	18.3
Pa.	1.31	2,169	2,856.3	9	13.7
Mich.	---	D			
Ohio	---	D			
Ill.	1.31	450	589	2	19.0
Wis.	---	D			
N.C.	---	D			
Fla.	---	D			
Ala.	---	D			
Tex.	---	D			
Ariz.	---	D			
Col.	---	D			
Calif.	---	D			

TABLE 2-13. SECTOR #40 HOUSEHOLD TEXTILES

Equivalent S.I.C. Sector(s): 239

Census of Manuf. (CM) Nat'l Purified Value of Shipments ($\$x10^6$): 5,887.8

National Sector Output per Worker ($\$x10^3$): 27.4

Nat'l Purif. Corr. Factor: 1.15 Scaling Factor: 1.25 1972 Price Index: 1.05

STATE	SEAS(modified) / CM(purified)	Purified CM Amt. ($\$x10^6$)	Modified SEAS Amt. ($\$x10^6$)	% of Nat'l CM Output	State O/W
Mass.	1.31	219	286.4	4	22.3
N.Y.	1.13	926	1,051.0	16	26.1
N.J.	1.33	311	412.6	5	23.9
Pa.	2.11	198	418.1	3	17.9
Mich.	.47	894	418.3	15	48.6
Ohio	.61	214	131.0	4	31.0
Ill.	1.03	281	290.5	5	28.4
Wis.	.96	48	46.0	<1	22.1
N.C.	.55	595	325.6	10	32.1
Fla.	1.11	54	60.0	<1	15.2
Ala.	.46	135	62.5	2	37.7
Tex.	1.51	138	208.8	2	22.2
Ariz.	.08	13	1.0	<1	15.7
Col.	3.42	13	44.5	<1	18.3
Calif.	1.20	368	441.1	6	21.2

TABLE 2-14. SECTOR #45 HOUSEHOLD FURNITURE

Equivalent S.I.C. Sector(s): 251

Census of Manuf. (CM) Nat'l Purified Value of Shipments ($\$x10^{6}$): 7,385.0

National Sector Output per Worker ($\$x10^{3}$): 23.35

Nat'l Purif. Corr. Factor: 1.00 Scaling Factor: .98 1972 Price Index: 1.02

STATE	SEAS(modified) / CM(purified)	Purified CM Amt. ($\$x10^{6}$)	Modified SEAS Amt. ($\$x10^{6}$)	% of Nat'l CM Output	State O/W
Mass.	1.39	172	238.5	2	23.9
N.Y.	1.42	320	455.2	4	22.9
N.J.	1.46	144	210.3	2	29.4
Pa.	1.33	300	398.9	4	23.6
Mich.	1.57	143	224.6	2	22.3
Ohio	.79	232	183.0	3	28.6
Ill.	1.33	336	446.6	4	28.5
Wis.	.97	97	94.6	1	22.0
N.C.	.84	1,374	1,157.5	18	21.7
Fla.	1.12	140	157.5	2	20.9
Ala.	.79	109	86.7	1	21.0
Tex.	1.02	247	251.9	3	23.7
Ariz.	.61	23	13.9	<1	20.9
Col.	1.40	28	39.4	<1	25.4
Calif.	.80	752	600.7	10	25.7

TABLE 2-15. SECTOR #51 PAPERBOARD PRODUCTS

Equivalent S.I.C. Sector(s): 265

Census of Manuf. (CM) Nat'l Purified Value of Shipments ($\$x10^{6}$): 8,148.1

National Sector Output per Worker ($\$x10^{3}$): 36.2

Nat'l Purif. Corr. Factor: 1.01 Scaling Factor: 1.04 1972 Price Index: 1.03

STATE	SEAS(modified) / CM(purified)	Purified CM Amt. ($\$x10^{6}$)	Modified SEAS Amt. ($\$x10^{6}$)	% of Nat'l CM Output	State O/W
Mass.	1.18	299	352.2	4	29.8
N.Y.	1.17	579	675.1	7	28.5
N.J.	1.37	498	681.4	6	32.6
Pa.	1.00	566	567.4	7	32.9
Mich.	1.07	385	410.8	5	41.4
Ohio	.98	590	579.5	7	36.3
Ill.	1.15	741	853.8	9	36.0
Wis.	.89	286	254.1	3	38.2
N.C.	.87	207	80.7	2	33.6
Fla.	.91	179	162.2	2	40.2
Ala.	1.10	42	46.3	<1	26.2
Tex.	.85	267	226.4	3	38.8
Ariz.	---	D	---		
Col.	---	D	---		
Calif.	.75	759	570.4	9	45.5

TABLE 2-16. SECTOR #52 NEWSPAPERS

Equivalent S.I.C. Sector(s): 271

Census of Manuf. (CM) Nat'l Purified Value of Shipments ($\$x10^6$): 8,200.0

National Sector Output per Worker ($\$x10^3$): 23.68

Nat'l Purif. Corr. Factor: .99 Scaling Factor: 1.03 1972 Price Index: 1.02

STATE	SEAS(modified) / CM(purified)	Purified CM Amt. ($\$x10^6$)	Modified SEAS Amt. ($\$x10^6$)	% of Nat'l CM Output	State O/W
Mass.	1.04	315	327.9	4	24.1
N.Y.	.98	921	903.5	11	28.1
N.J.	1.04	273	283.0	3	24.6
Pa.	1.02	496	504.0	6	21.8
Mich.	.67	355	239.5	4	27.0
Ohio	1.07	452	485.4	5	23.8
Ill.	1.11	545	605.4	7	24.7
Wis.	.73	170	123.4	2	20.2
N.C.	1.28	136	173.9	2	20.4
Fla.	.90	355	318.2	4	25.6
Ala.	.91	82	74.9	1	24.4
Tex.	1.08	403	434.7	5	23.1
Ariz.	.80	93	74.6	1	26.9
Col.	1.06	107	113.1	1	22.0
Calif.	.90	883	792.9	11	26.8

TABLE 2-17. SECTOR #53 PERIODICALS

Equivalent S.I.C. Sector(s): 272

Census of Manuf. (CM) Nat'l Purified Value of Shipments ($\$x10^6$): 3,584.7

National Sector Output per Worker ($\$x10^3$): 52.71

Nat'l Purif. Corr. Factor: 1.02 Scaling Factor: .79 1972 Price Index: 1.05

STATE	SEAS(modified) / CM(purified)	Purified CM Amt. ($\$x10^6$)	Modified SEAS Amt. ($\$x10^6$)	% of Nat'l CM Output	State O/W
Mass.	1.05	28	29.5	1	33.7
N.Y.	.62	1,776	1,105.7	50	67.5
N.J.	---	D	---		
Pa.	---	D	---		
Mich.	1.48	20	29.6	<1	33.0
Ohio	1.72	112	193.0	3	36.7
Ill.	1.09	325	354.7	9	47.6
Wis.	1.05	23	24.2	1	32.9
N.C.	---	D			
Fla.	.66	30	19.8	1	72.5
Ala.	---	NL			
Tex.	---	D			
Ariz.	---	NL			
Col.	2.63	19	50.0	<1	38.0
Calif.	.61	179	108.4	5	46.0

TABLE 2-18. SECTOR #54 BOOKS

Equivalent S.I.C. Sector(s): 273

Census of Manuf. (CM) Nat'l Purified Value of Shipments (\$x10^6): 4,140.4

National Sector Output per Worker (\$x10^3): 38.72

Nat'l Purif. Corr. Factor: 1.09 Scaling Factor: 1.14 1972 Price Index: 1.02

STATE	SEAS(modified) / CM(purified)	Purified CM Amt. (\$x10^6)	Modified SEAS Amt. (\$x10^6)	% of Nat'l CM Output	State O/W
Mass.	.77	278	214.3	6	35.1
N.Y.	.79	1,675	1,321.0	40	54.5
N.J.	2.53	133	334.8	3	38.1
Pa.	1.10	164	180.8	4	24.6
Mich.	.93	69	63.9	2	28.6
Ohio	1.06	207	220.2	5	41.3
Ill.	.99	571	563.3	14	58.2
Wis.	1.01	83	83.8	2	40.0
N.C.	1.10	16	17.7	<1	18.7
Fla.	.40	22	8.8	<1	22.2
Ala.	---	D			
Tex.	2.11	56	118.3	1	23.2
Ariz.	---	NL			
Col.	.90	19	17.0	<1	21.2
Calif.	1.16	161	186.2	4	29.6

TABLE 2-19. SECTOR #59 FERTILIZERS

Equivalent S.I.C. Sector(s): 2873, 2874, 2875

Census of Manuf. (CM) Nat'l Purified Value of Shipments (\$x10^6): 2,909.8

National Sector Output per Worker (\$x103): 77.85

Nat'l Purif. Corr. Factor: 1.05 Scaling Factor: 1.53 1972 Price Index: 1.01

STATE	SEAS(modified) / CM(purified)	Purified CM Amt. (\$x10^6)	Modified SEAS Amt. (\$x10^6)	% of Nat'l CM Output	State O/W
Mass.	---	NL			
N.Y.	1.21	21	24.3	1	65.3
N.J.	4.39	9	39.5	<1	43.5
Pa.	---	D			
Mich.	---	D			
Ohio	---	D			
Ill.	---	D			
Wis.	1.17	17	19.9	1	80.5
N.C.	.72	133	96.2	4	57.7
Fla.	1.03	476	492.0	16	74.6
Ala.	---	D			
Tex.	.61	200	122.1	7	86.4
Ariz.	---	NL			
Col.	---	NL			
Calif.	---	D			

TABLE 2-20. SECTOR #66 DRUGS

Equivalent S.I.C. Sector(s): 283

Census of Manuf. (CM) Nat'l Purified Value of Shipments ($\$x10^6$): 7,830.4

National Sector Output per Worker ($\$x10^3$): 61.73

Nat'l Purif. Corr. Factor: .98 Scaling Factor: 1.02 1972 Price Index: 1.01

STATE	SEAS(modified) / CM(purified)	Purified CM Amt. ($\$x10^6$)	Modified SEAS Amt. ($\$x10^6$)	% of Nat'l CM Output	State O/W
Mass.	1.99	28	55.7	1	36.3
N.Y.	.62	1,169	719.7	15	72.7
N.J.	.96	1,865	1,788.5	24	67.0
Pa.	.75	1,108	835.2	14	86.3
Mich.	1.29	466	599.2	7	65.2
Ohio	1.06	119	126.4	1	40.3
Ill.	1.32	539	712.2	7	45.8
Wis.	1.08	36	38.7	<1	37.0
N.C.	---	D	---		
Fla.	.31	65	20.4	1	60.0
Ala.	---	NL	---		
Tex.	1.50	50	75.1	1	36.4
Ariz.	---	D	---		
Col.	---	D	---		
Calif.	1.28	376	480.9	5	44.1

TABLE 2-21. SECTOR #67 CLEANING AND TOILET PRODUCTS

Equivalent S.I.C. Sector(s): 284

Census of Manuf. (CM) Nat'l Purified Value of Shipments ($\$x10^6$): 9,998.4

National Sector Output per Worker ($\$x10^3$): 87.5

Nat'l Purif. Corr. Factor: 1.02 Scaling Factor: 1.08 1972 Price Index: 1.00

STATE	SEAS(modified) / CM(purified)	Purified CM Amt. ($\$x10^6$)	Modified SEAS Amt. ($\$x10^6$)	% of Nat'l CM Output	State O/W
Mass.	.54	438	236.7	4	104.7
N.Y.	1.71	741	1,266.8	7	64.3
N.J.	1.07	1,542	1,653.4	15	73.1
Pa.	1.51	187	282.3	2	50.7
Mich.	---	D	---		
Ohio	1.29	1,044	1,345.7	10	146.3
Ill.	.77	1,414	1,089.7	14	112.7
Wis.	---	D	---		
N.C.	.69	115	79.5	1	53.4
Fla.	.45	58	26.3	<1	52.0
Ala.	---	NL	---		
Tex.	1.18	208	245.0	2	88.6
Ariz.	---	D			
Col.	.66	20	13.2	<1	100.5
Calif.	1.39	956	1,326.9	10	86.8

TABLE 2-22. SECTOR #69 PETROLEUM REFINING

Equivalent S.I.C. Sector(s): 291,299

Census of Manuf. (CM) Nat'l Purified Value of Shipments ($\$x10^6$): 28,163.4

National Sector Output per Worker ($\$x10^3$): 243.29

Nat'l Purif. Corr. Factor: 1.05 Scaling Factor: 1.01 1972 Price Index: 1.02

STATE	SEAS(modified) / CM(purified)	Purified CM Amt. ($\$x10^6$)	Modified SEAS Amt. ($\$x10^6$)	% of Nat'l CM Output	State O/W
Mass.	---	D			
N.Y.	---	D			
N.J.	---	D			
Pa.	.73	1,850	1,345.6	7	169.4
Mich.	---	D			
Ohio	.61	1,221	745.3	4	314.3
Ill.	1.41	1,869	2,632.2	7	240.5
Wis.	---	D			
N.C.	---	NL			
Fla.	---	NL			
Ala.	3.90	61	237.7	<1	193.3
Tex.	1.12	7,659	8,603.2	27	238.4
Ariz.	---	D			
Col.	---	D			
Calif.	.93	3,390	3,143.2	12	230.6

TABLE 2-23. SECTOR #76 FOOTWEAR

Equivalent S.I.C. Sector(s): 313,314

Census of Manuf. (CM) Nat'l Purified Value of Shipments ($\$x10^6$): 3,486.1

National Sector Output per Worker ($\$x10^3$): 18.8

Nat'l Purif. Corr. Factor: 1.00 Scaling Factor: .87 1972 Price Index: 1.07

STATE	SEAS(modified) / CM(purified)	Purified CM Amt. ($\$x10^6$)	Modified SEAS Amt. ($\$x10^6$)	% of Nat'l CM Output	State O/W
Mass.	1.50	413	617.6	12	19.0
N.Y.	1.00	230	229.7	7	30.1
N.J.	1.44	24	34.5	1	20
Pa.	1.18	326	384.1	9	16.7
Mich.	---	D			
Ohio	---	D			
Ill.	.77	139	106.6	4	20.7
Wis.	---	D			
N.C.	.49	65	32.0	2	21.7
Fla.	1.49	24	35.7	1	12.6
Ala.	---	D			
Tex.	.66	50	33.2	1	18.5
Ariz.	---	NL			
Col.	---	NL			
Calif.	.73	53	38.5	1	16.6

TABLE 2-24. SECTOR #81 CEMENT, CONCRETE, GYPSUM
Equivalent S.I.C. Sector(s): 324,327
Census of Manuf. (CM) Nat'l Purified Value of Shipments ($\$x10^6$): 9,398.0
National Sector Output per Worker ($\$x10^3$): 42.69
Nat'l Purif. Corr. Factor: .99 Scaling Factor: 1.04 1972 Price Index: 1.05

STATE	SEAS(modified) / CM(purified)	Purified CM Amt. ($\$x10^6$)	Modified SEAS Amt. ($\$x10^6$)	% of Nat'l CM Output	State O/W
Mass.	1.15	149	171.7	2	39.5
N.Y.	---	D			
N.J.	.94	256	239.6	3	43.9
Pa.	1.12	479	535.4	5	41.7
Mich.	.70	462	324.3	5	51.3
Ohio	1.16	469	543.2	5	41.6
Ill.	1.23	376	462.7	4	43.7
Wis.	1.21	175	211.3	2	38.5
N.C.	---	D			
Fla.	---	D			
Ala.	1.47	168	247.1	2	37.0
Tex.	.98	714	697.9	8	43.2
Ariz.	---	D			
Col.	---	D			
Calif.	.85	941	797.9	10	49.7

TABLE 2-25. SECTOR #84 COPPER
Equivalent S.I.C. Sector(s): 3331, 3341, 3351, 3362
Census of Manuf. (CM) Nat'l Purified Value of Shipments ($\$x10^6$): 12,734.6
National Sector Output per Worker ($\$x10^3$): 144.38
Nat'l Purif. Corr. Factor: 1.51 Scaling Factor: 2.01 1972 Price Index: 1.04

STATE	SEAS(modified) / CM(purified)	Purified CM Amt. ($\$x10^6$)	Modified SEAS Amt. ($\$x10^6$)	% of Nat'l CM Output	State O/W
Mass.	---	D			
N.Y.	---	D			
N.J.	---	D			
Pa.	1.23	489	602.7	4	53.5
Mich.	---	D			
Ohio	1.67	294	492.3	3	70.2
Ill.	---	D			
Wis.	---	D			
N.C.	---	D			
Fla.	---	NL			
Ala.	---	D			
Tex.	---	D			
Ariz.	---	NL			
Col.	---	NL			
Calif.	---	D			

TABLE 2-26. SECTOR #89 OTHER NONFERROUS ROLLING AND DRAWING

Equivalent S.I.C. Sector(s): 3356

Census of Manuf. (CM) Nat'l Purified Value of Shipments ($\$x10^6$): 1,197.7

National Sector Output per Worker ($\$x10^3$): 68.76

Nat'l Purif. Corr. Factor: .96 Scaling Factor: 1.22 1972 Price Index: 1.02

STATE	SEAS(modified) / CM(purified)	Purified CM Amt. ($\$x10^6$)	Modified SEAS Amt. ($\$x10^6$)	% of Nat'l CM Output	State O/W
Mass.	.87	108	93.6	9	113.0
N.Y.	.45	147	66.3	12	127.5
N.J.	.54	105	57.1	9	68.1
Pa.	.58	86	50.2	7	56.2
Mich.	1.12	53	59.3	4	68.7
Ohio	1.27	111	141.1	9	44.6
Ill.	---	D			
Wis.	---	NL			
N.C.	---	D			
Fla.	---	NL			
Ala.	---	D			
Tex.	---	D			
Ariz.	---	NL			
Col.	---	NL			
Calif.	---	D			

TABLE 2-27. SECTOR #94 PLUMBING AND HEATING EQUIPMENT

Equivalent S.I.C. Sector(s): 343

Census of Manuf. (CM) Nat'l Purified Value of Shipments ($\$x10^6$): 2,043.3

National Sector Output per Worker ($\$x10^3$): 35.2

Nat'l Purif. Corr. Factor: .96 Scaling Factor: 1.04 1972 Price Index: 1.03

STATE	SEAS(modified) / CM(purified)	Purified CM Amt. ($\$x10^6$)	Modified SEAS Amt. ($\$x10^6$)	% of Nat'l CM Output	State O/W
Mass.	.64	30	19.3	1	34.8
N.Y.	.96	131	126.4	6	33.2
N.J.	1.33	56	74.3	3	33.9
Pa.	.90	180	161.1	9	33.3
Mich.	1.02	95	97.2	5	39.5
Ohio	.99	203	200.1	10	39.0
Ill.	1.20	196	234.8	9	34.6
Wis.	1.07	132	141.3	6	32.7
N.C.	.12	29	3.4	1	33.2
Fla.	.52	9	4.7	<1	43.5
Ala.	---	D			
Tex.	---	D			
Ariz.	---	NL			
Col.	---	NL			
Calif.	---	D			

TABLE 2-28. SECTOR #95 STRUCTURAL METAL PRODUCTS

Equivalent S.I.C. Sector(s): 344

Census of Manuf. (CM) Nat'l Purified Value of Shipments ($\$x10^6$): 14,525.4

National Sector Output per Worker ($\$x10^3$): 35.94

Nat'l Purif. Corr. Factor: 1.02 Scaling Factor: 1.03 1972 Price Index: 1.04

STATE	SEAS(modified) / CM(purified)	Purified CM Amt. ($\$x10^6$)	Modified SEAS Amt. ($\$x10^6$)	% of Nat'l CM Output	State O/W
Mass.	1.12	264	296.9	2	32.4
N.Y.	1.20	829	991.7	6	31.3
N.J.	1.00	476	474.5	3	35.9
Pa.	1.18	1,237	1,457.9	8	36.0
Mich.	.95	582	551.9	4	36.8
Ohio	.89	1,278	1,136.5	9	38.9
Ill.	1.19	870	1,032.3	6	40.0
Wis.	1.31	356	467.9	2	34.2
N.C.	.79	322	254.4	2	41.6
Fla.	.67	556	374.5	4	34.5
Ala.	1.08	312	336.5	2	37.8
Tex.	1.14	1,009	1,150.5	7	33.9
Ariz.	1.13	93	105.4	1	33.7
Col.	1.65	97	159.9	1	32.8
Calif.	.95	1,195	1,131.1	8	37.6

TABLE 2-29. SECTOR #97 METAL STAMPING

Equivalent S.I.C. Sector(s): 3465, 3466, 3469

Census of Manuf. (CM) Nat'l Purified Value of Shipments ($\$x10^6$): 8,351.1

National Sector Output per Worker ($\$x10^3$): 37.2

Nat'l Purif. Corr. Factor: 1.00 Scaling Factor: 1.08 1972 Price Index: 1.05

STATE	SEAS(modified) / CM(purified)	Purified CM Amt. ($\$x10^6$)	Modified SEAS Amt. ($\$x10^6$)	% of Nat'l CM Output	State O/W
Mass.	---	D	---		
N.Y.	---	D	---		
N.J.	1.79	187	333.8	2	30.2
Pa.	---	D	---		
Mich.	.90	2,424	2,189.7	29	43.1
Ohio	.88	1,676	1,474.2	20	40.5
Ill.	1.08	998	1,081.7	12	40.1
Wis.	1.65	303	498.7	4	28.9
N.C.	---	D	---		
Fla.	---	D	---		
Ala.	---	D	---		
Tex.	3.02	17	51.3	<1	23.7
Ariz.	---	NL			
Col.	---	D			
Calif.	1.51	227	343.7	3	30.3

TABLE 2-30. SECTOR #98 CUTLERY, HANDTOOLS, HARDWARE

Equivalent S.I.C. Sector(s): 342

Census of Manuf. (CM) Nat'l Purified Value of Shipments ($\$x10^6$): 5,111.2

National Sector Output per Worker ($\$x10^3$): 31.9

Nat'l Purif. Corr. Factor: 1.00 Scaling Factor: .98 1972 Price Index: 1.03

STATE	SEAS(modified) CM(purified)	Purified CM Amt. ($\$x10^6$)	Modified SEAS Amt. ($\$x10^6$)	% of Nat'l CM Output	State O/W
Mass.	1.27	335	426.1	7	37.6
N.Y.	1.04	390	405.5	8	27.9
N.J.	.93	268	248.9	5	31.2
Pa.	1.33	193	257.2	4	25.4
Mich.	.70	731	509.1	14	43.0
Ohio	.84	672	563.0	13	41.7
Ill.	1.00	537	536.6	10	29.2
Wis.	.42	199	84.4	4	38.3
N.C.	.80	55	43.8	1	25.0
Fla.	.65	32	20.7	<1	24.6
Ala.	1.20	11	13.2	<1	15.7
Tex.	2.41	17	40.9	<1	28.3
Ariz.	---	NL	---		
Col.	.86	29	24.8	<1	26.4
Calif.	.90	415	371.9	8	29.9

TABLE 2-31. SECTOR #101 OTHER FABRICATED METAL PRODUCTS

Equivalent S.I.C. Sector(s): 347, 3493, 3497, 3499

Census of Manuf. (CM) Nat'l Purified Value of Shipments ($\$x10^6$): 4,288.3

National Sector Output per Worker ($\$x10^3$): 27.0

Nat'l Purif. Corr. Factor: ______ Scaling Factor: .85 1972 Price Index: 1.05

STATE	SEAS(modified) CM(purified)	Purified CM Amt. ($\$x10^6$)	Modified SEAS Amt. ($\$x10^6$)	% of Nat'l CM Output	State O/W
Mass.	1.09	91	99.4	2	20.1
N.Y.	---	D			
N.J.	1.22	202	245.5	5	28.5
Pa.	---	D			
Mich.	1.14	312	354.3	7	24.8
Ohio	---	D			
Ill.	.83	670	553.7	16	34.0
Wis.	2.67	64	170.9	1	20.5
N.C.	.94	28	26.4	<1	19.9
Fla.	1.07	49	52.6	<1	22.2
Ala.	---	D			
Tex.	---	D			
Ariz.	1.86	5	9.3	<1	16.0
Col.	1.12	22	24.7	<1	27.9
Calif.	---	D			

TABLE 2-32. SECTOR #103 FARM MACHINERY

Equivalent S.I.C. Sector(s): 352
Census of Manuf. (CM) Nat'l Purified Value of Shipments ($\$x10^6$): 5,501.7
National Sector Output per Worker ($\$x10^3$): 44.62
Nat'l Purif. Corr. Factor: .98 Scaling Factor: 1.23 1972 Price Index: 1.04

STATE	SEAS(modified) / CM(purified)	Purified CM Amt. ($\$x10^6$)	Modified SEAS Amt. ($\$x10^6$)	% of Nat'l CM Output	State O/W
Mass.	---	NL			
N.Y.	1.09	76	83.2	1	51.7
N.J.	---	D			
Pa.	---	D			
Mich.	.74	339	251.1	6	93.6
Ohio	1.02	254	260.2	5	45.4
Ill.	1.16	970	1,120.6	18	45.9
Wis.	.77	643	493.3	12	51.3
N.C.	1.29	60	77.6	1	27.7
Fla.	.59	25	14.7	<1	41.7
Ala.	---	D			
Tex.	1.77	59	104.5	1	31.4
Ariz.	---	D			
Col.	1.57	25	39.2	<1	28.8
Calif.	---	D			

TABLE 2-33. SECTOR #107 MACHINE TOOL, METAL FORMING

Equivalent S.I.C. Sector(s): 3542
Census of Manuf. (CM) Nat'l Purified Value of Shipments ($\$x10^6$): 709.7
National Sector Output per Worker ($\$x10^3$): 28.78
Nat'l Purif. Corr. Factor: 1.02 Scaling Factor: .92 1972 Price Index: 1.02

STATE	SEAS(modified) / CM(purified)	Purified CM Amt. ($\$x10^6$)	Modified SEAS Amt. ($\$x10^6$)	% of Nat'l CM Output	State O/W
Mass.	---	D			
N.Y.	1.08	63	68.1	9	26.8
N.J.	1.23	10	12.3	1	25.8
Pa.	1.15	34	39.1	5	27.4
Mich.	.74	88	65.3	12	33.2
Ohio	.70	161	112.0	23	30.4
Ill.	1.03	204	210.6	29	30.3
Wis.	---	NL			
N.C.	---	D			
Fla.	1.18	4	4.7	1	19.0
Ala.	---	NL			
Tex.	.72	5	3.6	1	24.0
Ariz.	---	NL			
Col.	---	D			
Calif.	---	D			

TABLE 2-34. SECTOR #113 INDUSTRIAL PATTERNS

Equivalent S.I.C. Sector(s): 3565, 3567, 3569

Census of Manuf. (CM) Nat'l Purified Value of Shipments ($\$x10^6$): 1842.4

National Sector Output per Worker ($\$x10^3$): 30.7

Nat'l Purif. Corr. Factor: 1.02 Scaling Factor: .92 1972 Price Index: 1.02

STATE	SEAS(modified) / CM(purified)	Purified CM Amt. ($\$x10^6$)	Modified SEAS Amt. ($\$x10^6$)	% of Nat'l CM Output	State O/W
Mass.	1.02	60	61.0	3	32.8
N.Y.	.90	126	113.7	7	31.9
N.J.	.53	124	65.4	7	32.9
Pa.	.69	224	153.9	12	31.0
Mich.	.99	192	190.1	10	31.3
Ohio	---	D			
Ill.	1.35	136	183.8	7	28.4
Wis.	---	D			
N.C.	.49	22	10.8	1	27.1
Fla.	---	D			
Ala.	---	D			
Tex.	1.12	22	24.6	1	24.9
Ariz.	---	NL			
Col.	---	NL			
Calif.	---	D			

TABLE 2-35. SECTOR #118 ELECTRONIC MEASURING INSTRUMENTS

Equivalent S.I.C. Sector(s): 3825

Census of Manuf. (CM) Nat'l Purified Value of Shipments ($\$x10^6$): 1,408.6

National Sector Output per Worker ($\$x10^3$): 28.2

Nat'l Purif. Corr. Factor: .91 Scaling Factor: 1.08 1972 Price Index: 1.01

STATE	SEAS(modified) / CM(purified)	Purified CM Amt. ($\$x10^6$)	Modified SEAS Amt. ($\$x10^6$)	% of Nat'l CM Output	State O/W
Mass.	.69	86	59.2	6	27.6
N.Y.	1.03	110	113.4	8	35.6
N.J.	1.81	53	95.8	4	24.2
Pa.	1.16	44	51.1	3	24.0
Mich.	.62	25	15.6	2	24.5
Ohio	1.17	63	73.5	5	24.6
Ill.	.96	132	126.4	10	30.2
Wis.	---	D			
N.C.	---	D			
Fla.	3.44	5	17.2	<1	12.5
Ala.	---	NL			
Tex.	.69	34	23.3	2	28.5
Ariz.	---	NL			
Col.	---	D			
Calif.	---	D			

TABLE 2-36. SECTOR #119 TRANSFORMERS AND SWITCHES

Equivalent S.I.C. Sector(s): 361

Census of Manuf. (CM) Nat'l Purified Value of Shipments ($\$x10^6$): 3,605.2

National Sector Output per Worker ($\$x10^3$): 31.23

Nat'l Purif. Corr. Factor: 1.01 Scaling Factor: 1.00 1972 Price Index: 0.99

STATE	SEAS(modified) / CM(purified)	Purified CM Amt. ($\$x10^6$)	Modified SEAS Amt. ($\$x10^6$)	% of Nat'l CM Output	State O/W
Mass.	.33	187	62.5	5	25.3
N.Y.	1.56	79	123.4	2	30.0
N.J.	1.23	142	175.3	4	25.2
Pa.	1.41	799	1,122.6	22	30.0
Mich.	1.38	37	50.9	1	33.6
Ohio	.95	154	146.8	4	32.3
Ill.	1.31	315	412.1	9	26.2
Wis.	.51	192	97.6	5	40.4
N.C.	1.00	103	103.5	2	30.0
Fla.	---	D			
Ala.	---	D			
Tex.	.90	60	54.1	2	39.3
Ariz.	---	D			
Col.	.62	6	3.7	<1	30.0
Calif.	1.00	194	193.8	5	32.5

TABLE 2-37. SECTOR #123 HOUSEHOLD APPLIANCES

Equivalent S.I.C. Sector(s): 363

Census of Manuf. (CM) Nat'l Purified Value of Shipments ($\$x10^6$): 6973.6

National Sector Output per Worker ($\$x10^3$): 42.72

Nat'l Purif. Corr. Factor: 1.01 Scaling Factor: .92 1972 Price Index: 1.00

STATE	SEAS(modified) / CM(purified)	Purified CM Amt. ($\$x10^6$)	Modified SEAS Amt. ($\$x10^6$)	% of Nat'l CM Output	State O/W
Mass.	.55	75	41.5	1	38.9
N.Y.	---	D			
N.J.	.85	103	87.5	1	24.3
Pa.	---	D			
Mich.	---	D			
Ohio	1.38	1,165	1,602.6	17	49.5
Ill.	1.03	863	891.5	12	37.0
Wis.	.98	282	276.3	4	35.3
N.C.	.78	171	132.8	2	36.7
Fla.	---	D			
Ala.	.27	19	5.1	<1	38.0
Tex.	.55	43	23.6	1	30.7
Ariz.	---	NL			
Col.	---	D			
Calif.	.75	256	191.9	4	41.5

TABLE 2-38. SECTOR #127 COMMUNICATION EQUIPMENT

Equivalent S.I.C. Sector(s): 366

Census of Manuf. (CM) Nat'l Purified Value of Shipments ($\$x10^6$): 13,353.6

National Sector Output per Worker ($\$x10^3$): 30.13

Nat'l Purif. Corr. Factor: .98 Scaling Factor: .87 1972 Price Index: 1.00

STATE	SEAS(modified) / CM(purified)	Purified CM Amt. ($\$x10^6$)	Modified SEAS Amt. ($\$x10^6$)	% of Nat'l CM Output	State O/W
Mass.	1.60	820	1,313.7	6	29.2
N.Y.	1.06	1,231	1,303.7	9	27.0
N.J.	1.08	1,178	1,270.2	9	30.3
Pa.	.80	466	373.5	4	31.7
Mich.	.54	101	54.2	1	42.9
Ohio	1.13	387	437.8	3	30.6
Ill.	.98	1,442	1,411.2	11	31.1
Wis.	1.39	83	115.6	1	25.0
N.C.	---	D			
Fla.	1.26	425	537.2	3	26.1
Ala.	.47	177	83.1	1	27.5
Tex.	---	D			
Ariz.	---	D			
Col.	---	D			
Calif.	.99	2,400	2,383.8	18	32.4

TABLE 2-39. SECTOR #128 ELECTRONIC COMPONENTS

Equivalent S.I.C. Sector(s): 367

Census of Manuf. (CM) Nat'l Purified Value of Shipments ($\$x10^6$): 9038.0

National Sector Output per Worker ($\$x10^3$): 26.35

Nat'l Purif. Corr. Factor: 1.02 Scaling Factor: 1.00 1972 Price Index: 1.00

STATE	SEAS(modified) / CM(purified)	Purified CM Amt. ($\$x10^6$)	Modified SEAS Amt. ($\$x10^6$)	% of Nat'l CM Output	State O/W
Mass.	1.86	396	735.6	4	21.8
N.Y.	.50	2,046	1,020.0	23	41.7
N.J.	2.24	361	809.4	4	19.7
Pa.	1.04	882	914.1	10	27.5
Mich.	1.32	53	69.8	<1	20.8
Ohio	.93	263	245.8	3	32.2
Ill.	1.29	580	748.2	6	22.6
Wis.	1.66	69	114.3	<1	21.9
N.C.	---	D			
Fla.	.74	219	162.1	2	23.4
Ala.	---	D			
Tex.	---	D			
Ariz.	.89	327	292.2	4	19.8
Col.	3.52	14	49.3	<1	14.0
Calif.	1.16	1,365	1,585.3	15	25.2

CHAPTER 3

AN ASSESSMENT OF LOCATION QUOTIENT METHODS OF REGIONALIZING THE NATIONAL INPUT-OUTPUT MODEL

I. INTRODUCTION

The depiction of a regional economy by an input-output table is useful in both a descriptive and predictive sense. The interrelationships of all parts of a regional economy are comprehensively detailed in such a model, and its structure permits detailed analysis of the workings of that economy. However, data collection for such a detailed table requires expensive surveys, and construction of the table often requires such a lengthy period that it is often obsolete before it is published.

Various short-cut methods have been proposed and tested for deriving a regional table from a given national table, using only a limited amount of regional data. Such an approach is attractive, as it is relatively simple, inexpensive, and quick, and requires no extensive data gathering. However, such techniques have yet to be proven effective.

Past attempts to regionalize national input-output tables have produced less than satisfactory results. The requirement of having available a pre-existing ("true") regional table for verification of the estimate has imposed constraints on the degree of disaggregation of industry sectors, as the sectors of the estimating table must correspond to those of the "true" table in order for meaningful comparisons of results to be made. As a result, typical sectoring schemes in previous studies have been forced to use less than 30 sectors in order to be compatible with the "true" regional table while national tables are available at a 367-sector level. The aggregation necessary to reconcile the sectors of the national table with those of the existing regional table introduces a wide industry (or product) mix within most sectors. This, together with the inherent limitations in the value of highly aggregated input-output tables, disturbs the premises necessary for application and examination of short-cut regionalization techniques for national tables.

The present study is intended to evaluate the method of location quotients (measures of the relative importance of a regional industry) as a means of modifying a national input-output table to reflect the structure of a regional economy. If the application of such a method is indeed verified as a valid means of estimating the interindustry structure of a regional economy, requiring a minimum of data and avoiding expensive survey procedures, then an alternative may be available for the existing method of regional projection of the SEAS system. As opposed to the current top-down structure involving national projections allocated to regions by shares based on employment, the easy availability of regional input-output tables would allow accurate regional projections to then be aggregated to a national level. Alternatively, the current method of national projections could be retained, while specific projections for those regions under consideration could be derived using those regional tables.

We will first survey quotient methods previously used to regionalize national input-output tables.[1] Following this, we describe our ideas for improving the means of evaluating the effectiveness of one method in particular, that of simple location quotients. The next section will describe our procedure for examining this method. The means of measuring the technique's simulation ability will be detailed, and finally our results and conclusions will be presented.

II. QUOTIENT METHODS: A BRIEF SURVEY

Morrison and Smith (1974) describe several quotient methods to regionalize national input-output tables which use the principle that $a_{ij} = q_{ij} A_{ij}$, where a_{ij} = the regional trade coefficient from industry i to industry j (i.e. the ratio of i's input to j divided by j's total output), A_{ij} = the national technical coefficient (i.e. the national ratio), and the reduction quotient q_{ij} is less than or equal to 1. Exports are obtained as residuals:

$$e_i^R = x_i^R - \sum_j a_{ij} x_j^R - y_{if}^R, \quad .$$

where y_{if}^R is final demand for industry i, x is output, the subscripts i and j are sectors, and the superscript R denotes the region.[2] Imports of product i are computed as the amounts necessary to satisfy production requirements not met by regional industry i, assuming national production technology holds in region R:

$$m_{ij}^R = A_{ij} x_j^R - a_{ij} x_j^R.$$

Simple Location Quotient

The simple location quotient (SLQ) measures the relative importance of the industry in the region compared to its importance in the nation:

$$LQ_i = (N_i^R / N^R) \;/\; (N_i / N).$$

---

[1] A summary of other previous studies of regionalization techniques is contained in Section VIII of this chapter.

[2] Equations and notation in the following pages will always be assumed to refer to region R, and the superscript will be omitted except where confusion may otherwise result.

Here, N represents output, employment, or some other measure of economic importance. If $LQ_i \geq 1$, then the region is assumed to be self-sufficient in industry i, and $r_{ij} = a_{ij}$, the national coefficient. If $LQ_i < 1$, then local requirements are not satisfied locally, and $r_{ij} = LQ_i\, a_{ij}$.

Purchases-Only Location Quotient

The use of a "Purchases-Only Location Quotient" (POLQ) was suggested by Tiebout (Consad, 1967, pp. 3.10-3.13) and differs from the simple location quotient in that the summation of total output (or employment) includes only those sectors purchasing from sector i. Consistent results were obtained when used in the Consad study, and the POLQ ranked directly below the SLQ in all five of the tests of Morrison and Smith, who note only a marginal difference in results and point out that Schaffer and Chu (1969) do not distinguish between the two in their discussion of results.

Consad Alternatives

In an appendix to the Consad study, two alternate methods were examined and rejected in favor of POLQ (Consad, 1967, pp. F.1-F.6). In the first method,

$$LQ_i^R = (N_i^R/\Sigma_j N_j^R A_{ij}) \;/\; (N_j/\Sigma_j N_j A_{ij})$$

where j refers to consuming industries, and the national technical coefficients are incorporated into the definition to weight what would otherwise be the Purchases-Only Location Quotient. In general, results obtained were not significantly different from the POLQ method, with both positive and negative variations in individual elements. In all cases, this alternate method resulted in slightly higher values (1% to 5%) than POLQ. Consad concluded that, as results were so similar and intra-state self-sufficiency was already considered overstated, the POLQ results were to be used.

The second alternative considered by Consad, based on interstate imports, used a location quotient of the form

$$LQ_i^R = N_i^R/(N_i^R + I_i^R)$$

where I_i^R = imports of industry i into region R. In recognition of the extensive data requirements, the direct estimate of the Washington model's imports (considered of high accuracy) and imports imputed from Census of Transportation data were compared. Large discrepancies existed between the two sets of data in all cases, and as a result no further consideration of this alternative was made.

Cross-Industry Quotient (Morrison and Smith, 1974)

In order to include the relative size of the purchasing industry as well as that of the producing sector, the Cross-Industry Quotient (CIQ) compares the proportion of national output of producing sector i to that of purchasing sector j:

$$CIQ_{ij} = (N_i^R/N_i) \; / \; (N_j^R/N_j).$$

(Hence $CIQ_{ij} = LQ_i/LQ_j$.) This, in a manner similar to those above, is used to adjust the national coefficient.

Logarithmic Cross-Quotient

Round, using a Logarithmic Cross-Quotient (RND), takes into account the relative sizes of region and nation while retaining the properties of the CIQ:

$$RND_{ij} = LQ_i/\log_2(1 + LQ_j)$$

applied in the manner of the other quotients (Morrison and Smith, 1974).

Round's Welsh Model

Round (1972), developing his Welsh model with two regions and stressing the presence of a set of clearly defined assumptions (inputs and outputs of regional industries to and from national industries are in the same proportion as those of the national industries, and as far as possible, intraregional trade takes precedence over exporting), derives quotients for regions R_1 and R_2 similar to the cross-industry quotient, but where x = total intermediate output and Y = total intermediate input,

$$\alpha_{ij} = (X_i^{R1}/X_i) \; / \; (Y_j^{R1}/Y_j)$$

and

$$\beta_{ij} = (X_i^{R2}/X_i) \; / \; (Y_j^{R2}/Y_j).$$

Round stresses what he calls three essential points: that the Welsh model is as much a quotient technique as the others, that it has appeal as being based on well-defined assumptions lacking in other methods, and that a trading coefficient in one region defines the corresponding one in the other region and the model is internally consistent. However, Round closes by stating that " [i]t is difficult to believe that the highly complex pattern of interregional trade can be adequately described by a set of such extremely simple constructs. It is clear from the earlier tests of the Welsh model and more recent tests, as yet unpublished, that a great deal more research is yet to be carried out before one is in a position to assess the prospects of utilizing non-survey techniques in regional input-output analysis" (Round, 1972, p. 9).

Cross-Industry Quotient Modifications

Morrison and Smith (1974), noting that $CIQ_{ij} = LQ_i/LQ_j$ implies that intrasectoral shipments of commodity i (i.e., where i = j) have $CIQ_{ij} = 1$, and therefore that sector i is fully supplied locally, decided to substitute the simple location quotient for the elements of the CIQ_{ij} along the main diagonals to remove this problem in both cross-industry quotient techniques. These fifth and sixth quotient methods examined by Morrison and Smith were called the modified cross-industry quotient (CMOD) and modified logarithmic cross-quotient (RMOD).

Summary of Location Quotients

Isard (1960, pp. 123-126), succintly summarizes the characteristics of location quotients. Their advantages include simplicity and the fact that they can be based on readily available data. They possess four major disadvantages, however. Tastes and expenditure patterns of households of the same type and income differ over regions. For example, little fuel is used in the South, and much in the North, so that a location quotient of 1 for fuel in a Southern state would imply it exports, while in the North the state would be an importer. Income levels also differ over regions; so do production practices and industrial mixes. These factors all affect the location quotient, which is based on the assumption that all factors of an industry are in the same relative proportion to its output over all regions.

We have used the simple Location Quotient Method in our present study. A review of experience of other analysts with these and other methods of regionalizing national input-output models appears in the appendix to this chapter.

III. APPROACH TO ASSESSMENT OF THE LOCATION QUOTIENT METHOD

Previous evaluations of short-cut methods of regionalizing national input-output tables have concluded that the results may be of some help in estimating certain industry sectors, but that at this time the technique cannot be used reliably as a basis for regional studies. Iterative approximating techniques may be more intuitively satisfying than other methods, but the cost of their implementation would probably be prohibitive for repeated use in a system such as SEAS as a means of generating tables for all regions in the nation. Further, some iterative techniques require more data which must be obtained by survey methods, adding to the cost. Also, iterative techniques (e.g., Schaffer and Chu's R-I-O-T) have not necessarily produced better results than simpler techniques. In fact, location quotient methods have been shown generally to be among the better methods of estimation. Schaffer and Chu found them to be superior to their iterative procedure and a pool technique. Morrison and Smith's comparative tests for eight methods found only the semi-survey-based RAS iterative technique consistently better than the simple location quotient method. Because of the relatively good showing of location quotient methods in comparison with others, coupled with the simplicity and minimal data requirements of quotient models, it was decided that for our needs these methods provided the most promise for further examination.

In reviewing previous studies of location quotients, several handicaps to testing of the method were observed. The difficulties and their potential remedies are described below.

Computation of Location Quotients: Employment vs. Output

First, because of the problems of data availability, employment is generally considered as the surrogate measure of "importance" of an industry, rather than output. (Although Schaffer and Chu did use output in computing location quotients, they were hampered by other data problems which will be discussed shortly.) Intuitively, in the estimation of a regional economic input-output model, gross output by sector would seem to give a better measure of an industry's importance than would employment.

Compatibility of National and Regional Input-Output Tables Used

Secondly, the verification of the validity of any estimating method requires a valid data base (here, a national table) as a starting point, and a valid true depiction of what is being modeled, to compare with the estimate. This requirement has also proven to be a handicap to previous studies of regionalization methods. In particular, it means that the national table and the true regional table must be comparable in terms of time frame and sector classification and aggregation. Unless the estimated tables derived from the national table are comparable to the true regional table in these terms, comparison of estimated and true results is impossible. This problem has been a significant one in previous studies, in several aspects. The "true" regional table itself is based on survey methods, and so is itself actually only an estimate, containing a certain amount of error. Similar criticism may be made of the national input-output table used to begin the process. Also, the absence of a standard industry classification scheme for regional tables has resulted in a variety of sectoring breakdowns not comparable with national tables. Because of this, in order to compare corresponding compatible sectors of the "true" and estimated tables, aggregation of both the "true" and national tables are required before estimation begins. This results in an often drastic reduction in the number of sectors in the table, introducing a large product mix within a sector with all its attendant problems in input-output applications. Further, a table with a small number of highly aggregated sectors is of limited use to regional analysts. Finally, a "true" regional table must be found such that the economy it describes is of the same point in time as the national table on which the estimate is based. (Schaffer and Chu used the 1958 national table to estimate the 1963 survey-based table of Washington state.) Several of these problems were encountered in a previous study (Walderhaug, 1971) attempting to adjust the 1963 national table to conform to that year's 27-sector Washington table strictly by trying to achieve conformance between sector definitions and conventions. Walderhaug judged that the coefficients of over half the sectors of the adjusted table (14 of the 27) differed from the survey-based table due to technology and product mix, definitions and conventions, and questionable state sample data.

These problems have hampered previous studies, but suggest that if they could be overcome, the location quotient method could be tested more fully on its merits, rather than be subject to the limitations of structural data problems.

In fact, previous studies have recognized this, and suggested further investigation and elimination of some of these handicaps:

> In themselves the nonsurvey methods may prove useful supplements to survey studies...
>
> But it seems that, at the moment, there is still no acceptable substitute for a good survey-based study. This conclusion, however, is founded on a very limited test. We deliberately avoided any but grossly mechanical computations, and no systematic correction procedures are available yet. Further, our tests have been based on an aggregation procedure which completely ignores the product mix in Washington. Care in aggregation might provide quite different results -- it is surprising to us that our techniques yielded acceptable regional production coefficients for over 25 per cent of the industry categories.
>
> Further tests against other survey-based tables may provide evidence of consistent differences between our estimates and survey results. These differences might be used as adjusting factors to render our estimates more reliable and to produce at least pedagogically acceptable models of regional economies.
> (Schaffer and Chu, 1969, p. 96)

Measurement of the Accuracy of the Estimated Table

One additional problem also existed in previous attempts to evaluate regionalization methods. The measures of comparison between actual and estimated tables used in such studies have been numerous, but no single measure has been used by all. As a result, there is no common means to compare the results of all previous studies, and many of the measures used have little meaning on a cardinal scale. As no ideal comparative statistic appears to have been found, the best approach appears to have been followed by Morrison and Smith (1974) who used a variety of alternative statistics to rank methods (although publishing few numerical results).

Use of a Control Model to Measure the Improvement Attributable to the Technique

Another problem, pointed out and dealt with by McMenamin and Haring, was the neglect of testing whether the estimated table performed better than would a simple naive model. In a sense, this can allow a crude method of cost-effectiveness analysis, and would further demonstrate in real terms any advantage of estimating tables using location quotients. McMenamin and Haring (1974), using the RAS technique to estimate the 1967 Washington table from the 1963 Washington table, took the 1963 table as the naive model. In our study, the national table, unmodified by location quotients, was used as the naive model.

Modifications to the Simple Location Quotient

Modifications to the simple location quotient were presented in Section II of this chapter. However, previous studies of their applications to national tables have generally proven them to be inferior to the simple location quotient. The "purchases-only location quotient" which sums only over purchasing industries for that sector was reported to give "consistent results" by Consad (1967), and ranked immediately below the simple quotient in all of Morrison and Smith's tests (1974). As a result, no potential benefit was seen in testing this modification. The "cross-industry location quotient" (which takes into account the relative importance of both the selling and purchasing industries) provided Schaffer and Chu with results of comparable quality to those of the simple location quotient and best estimated imports and exports. However, Morrison and Smith's comparisons of regionalization methods ranked the cross-industry technique last or next-to-last in all tests. Again, due to these pessimistic results, we felt that further investigation of this modification was not a promising avenue to pursue. RND, CMOD, and LMOD similarly ranked low in Morrison and Smith's study.

Summary

So, based on our preliminary examination of the results and problems of previous studies, we decided to evaluate the method of regionalizing national input-output tables using simple location quotients in a comprehensive and systematic manner, while working to remove as many of the obstacles encountered by previous researchers as possible. Summarized, our goals were to test the simple location quotient technique with:

1) The least sector aggregation possible;

2) The most compatible sectoring scheme possible between the national table (and therefore the estimated table) and regional table;

3) The use of a large variety of statistical methods of comparison;

4) A comparative analysis of the advantages of the added information of output by sector over that of employment by sector, and of both of these over a naive model of unaltered national coefficients;

5) An analysis of the ability of the location quotient method to estimate a "true" regional input-output table, given the efforts described above.

IV. PROCEDURES USED TO ESTIMATE THE 1967 WASHINGTON STATE INPUT-OUTPUT TABLE

In order to examine the questions raised in the previous section on the usefulness of the location quotient method, a set of estimates of the 1967 Washington state input-output tables were developed. Examination was also made of several alternatives to this simple application of location quotients to the national direct coefficients matrix. Specifically, the tables were adjusted for internal consistency (non-negative exports). A detailed description of the procedure is presented in this section. The analysis of the results is described in sections V, VI, and VII.

Reconciliation of National and Washington Tables

The survey-based input-output table for the state of Washington in 1967[3] was assumed to be the "true" table as the basis for comparison. The 1967 national table of the Bureau of Economic Analysis[4] was used as the given data base for application of the location quotient method. Location quotients were to be applied to the national table, and the results compared to those directly obtainable from the "true" table.

---

[3] W. B. Beyers, et al., Input-Output Tables for the Washington Economy 1967, Seattle: Graduate School of Business Administration, University of Washington, 1970. The Washington intra-regional transactions table with imports as a separate row was used, rather than the table incorporating imports in the interindustry sales matrix. First, the rationale of applying location quotients to national tables is to reflect the level of regional self-sufficiency in supplying goods within a region where national technology is assumed. As such, strictly intra-regional transactions are the appropriate elements of the table. Second, it was felt that intra-regional relationships are of more importance to regional users than strict technological relationships are. Third, Czamanski and Malizia (1969) found in their test of the RAS method that the intra-regional table could be estimated more accurately than the one incorporating imports.

[4] 367-sector national tables are published in Social and Economic Statistics Administration/Bureau of Economic Analysis, United States Department of Commerce. Input-Output Structure of the US Economy: 1967, v. 1-3, Washington: United States Government Printing Office, 1974; magnetic tape is available from Bureau of Economic Analysis.

85-sector tables are also available on tape, and are published in Interindustry Economics Division, Bureau of Economic Analysis, United States Department of Commerce, "The Input-Output Structure of the U.S. Economy: 1967," Survey of Current Business, LIV, no. 2, (Feb. 1974) pp. 24-56.

The need to directly compare the 27- or 52-sector Washington table with the 85- or 367-sector national table while preserving the most disaggregated sectoring scheme possible led to the choice of a 50-sector classification scheme. The detail of the 367-sector U.S. table permitted very close reconciliation with Washington's 52 sectors, through aggregation (Table 3-1). Only three Washington sectors, fishing, forestry, and miscellaneous agricultural products, were forced to combine (ironically due to their aggregation within the U.S. table). While business and personal services sectors also could not be exactly reconciled, they were left as separate industries due to their individual importance within the economy. Both Washington and U.S. interindustry transactions tables were aggregated to this 50-sector configuration by pre- and post-multiplication by grouping matrices.[5]

---

[4] (continued)
Definitions, conventions and industry descriptions are contained in Social and Economic Statistics Administration/Bureau of Economic Analysis, Interindustry Economics Division, United States Department of Commerce, Definitions and Conventions of the 1967 Input-Output Study, October, 1974.

[5] One convention of the national table is its treatment of secondary products in primary producing sectors. Where this occurs, a transfer is created, depicting a flow of the secondary product from the producing industry into the sector in which it is primary, and counting it as an interindustry transaction. This transferred product is then distributed to its consuming industries from its new sector. In the aggregation process, any such transfers between industries which are to be combined must be subtracted from interindustry flow, as such a transfer would become intra-industry. No such transfers are performed in the Washington State tables.

Other conventions in the national tables should also be noted. Government sectors are treated as separate sectors in the national table, while incorporated in the corresponding industrial sector by the Washington table. Only for the government sectors listed separately by function could proper national aggregation into the corresponding industry sector be performed.

Competitive imports are transferred as fictitious purchases from the imports row to the domestic industry, then distributed by that industry to purchases.

See Beyers, et al., 1970, pp. 7-8 for Washington table's conventions, and U.S. Department of Commerce, Definitions for national table's conventions.

The national table's "special," "dummy," and unclassified government industries which had no Washington state counterpart were not included in any computations except total gross output by sector. Much of their sales were direct to final demand, so that their absence had little influence on the derivation of the national and estimated direct coefficients matrices.

TABLE 3-1. 1967 WASHINGTON 50-SECTOR INPUT-OUTPUT TABLE

Sectoring Scheme Used	Washington Sector No. (1-52)	Washington Sector Name	U.S. Sector No. (1-85)	U.S. Sector Name	S.I.C. No.[a]
1	1	Field crops, incl. seeds	2.01-2.03, 2.06	(Various field crops)	011,0142//011, pts. 014
2	2	Vegetable, fruits, & nut crops	2.04-2.05	Fruits & tree nuts, vegetables, sugar & misc. crops	012//012, pts. 014, pt. 0119
3	3	Livestock & livestock products	1	Livestock & livestock products	013,0142, 0144, 072// 013,pt.014, 0193,pt.0729
4	4	Special & Misc. products	2.07	(Forest, greenhouse, nursery products)	019,02,073,074// 0192,pt.014
	5	Fishing	3	(Forestry & fishery products)	09//074,081,082, 084,086,091
	15	Forestry	4	(Agricultural, forestry & fishery services)	08//071,0723,073 pt.0729,085,098
5	6	Meat products	14.01	Meat products	201
6	7	Dairy products	14.02-14.06	(Various dairy products)	202
7	8	Canning and preserving	14.07-14.13	(Various canning and preserving)	203
8	9	Grain mill products	14.14-14.17	(Various canning and preserving)	204
9	10	Beverage industries	14.21-14.23	(Various beverages)	208

Table 3-1. 1967 Washington 50-Sector Input-Output Table (Page two)

Sectoring Scheme Used	Washington Sector No. (1-52)	Washington Sector Name	U.S. Sector No. (1-85)	U.S. Sector Name	S.I.C. No.[a]
10	11	Other foods	14.18-14.20 14.24-14.32	(Various other foods)	205-7,209
11	12	Textile mill products	16,17 18.01-18.03	(Various textile mill products)	22
12	13	Apparel	18.04,19	(Various apparel)	23
13	14	Mining	5-10	(Various types of mining)	10-14
14	16	Logging	20.01	Logging camps & logging contractors	241
15	17	Sawmills	20.02	Sawmills & Planing Mills, General	2421
16	18	Veneer and Plywood	20.06	Veneer & plywood	2432
17	19	Misc. wood products	20.03-.05, 20.07,20.09 21	(Various lumber and wood products) wooden containers	2426,2429,2431, 2433,249 244
18	20	Furniture & fixtures	22 23	Household furniture Other furniture & fixtures	25
19	21	Pulpmills	24.01	Pulpmills	261
20	22	Papermills	24.02	Papermills	262
21	23	Misc. paper & paperboard mills	24.03-24.07 25	(Various paper & allied products) Paperboard containers and boxes	263-266

Table 3-1. 1967 Washington 50-Sector Input-Output Table (Page three)

Sectoring Scheme Used	Washington Sector No. (1-52)	Washington Sector Name	U.S. Sector No. (1-85)	U.S. Sector Name	S.I.C. No.[a]
22	24	Printing, publishing & allied Indus.	26	Printing and publishing	27
23	25	Indus. Inorg. & Org. chemicals	27.01	Indus. Inorg. & Org. chemicals	281
24	26	Other chemicals	27.02-27.04 28-30	(Various chemicals & products)	282-289
25	27	Petrol. Refin. & related Indus.	31	Petroleum Refin. & related Indus.	29
26	28	Glass & cut stone products	35 36.15-36.22	Glass & glass products (Various stone & clay	321-3 328-9
27	29	Cement, clay, concrete, gypsum & plaster products	36.01-36.14	(Various stone & clay products)	324-7
28	30	Iron and steel rolling & finishing mills, foundries & forging	37	Primary iron & steel Manuf.	331-332,3391, 3399
29	31	Nonferrous metals, exc. aluminum	38.01-.03, .05-.07, .09-.10, .12-.14	(Nonferrous metals, ex. aluminum)	3331-3,3339, 3341,3351,3356, 3357,3362,3369, 3392
30	32	Aluminum	38.04 38.08,38.11	(Various aluminum products)	3334,3352, 3361
31	33	"Heavy" fabricated metal products	40.04-40.09	(Various heavy metal products)	344
32	34	"Light" fabricated metal products	39,40.01-.03 41,42	(Various light metal products)	341-343,345-349

Table 3-1. 1967 Washington 50-Sector Input-Output Table (Page four)

Sectoring Scheme Used	Washington Sector No. (1-52)	Washington Sector Name	U.S. Sector No. (1-85)	U.S. Sector Name	S.I.C. No.[a]
33	35	Nonelectrical motive & moving equipment	43-46	(Various engines & machinery)	351,352,353
34	36	Machine tools & shops	47	Metalworking machinery & equip.	354
			50	Machine shop products	359
35	37	Industrial processing equip.	48,49,51,52	(Various industrial processing equip.)	355,356,357,358
36	38	Electrical machinery, equipment & supplies	53-58	(Various electrical equip.)	36
37	39	Aerospace	60	Aircraft & parts	372
38	40	Motor vehicles & equip., railroad cars	59 61.03-61.07	Motor vehicles & equip. (Various other trans. equip.)	371,374, 375,379
39	41	Shipbuilding & repairing	61.01-61.02	(Ship -& boat-building & repairing)	373
40	42	Other manufacturing	13,15,32-34, 62-64	(Various other manuf.)	19,21,30,31, 38,39
41	43	Transportation services	65, (state & local govt. 79.01)	Transportation & warehousing	40-47
42	44	Gas, water, sanitary services	68, (U.S. govt. 78.02--state & local govt. 79.02)	Electric, gas, water, & sanitary services	49

Table 3-1. 1967 Washington 50-Sector Input-Output Table (Page five)

Sectoring Scheme Used	Washington Sector No. (1-52)	Washington Sector Name	U.S. Sector No. (1-85)	U.S. Sector Name	S.I.C. No.[a]
43	45	Communications	66 67	Communications exc. radio & TV Radio & TV broadcasting	48
44	46	Construction	11 12	New construction Maintenance & repair const.	15,16,17
45	47	Wholesale, retail trade	69	Wholesale & retail trade	50,52-59
46	48	Finance	70.01-70.03	(Various banking businesses)	60-62, 67
47	49	Insurance	70.04-70.05	Insurance carriers, agents & brokers	63-64
48	50	Real estate	71	Real estate & rental	65,66,exclude 656
49	51	Business service	73	Business services	73,81,861, 862,89//[b]
50	52	Personal service	72,75-77	(Various personal services)	70-72,75-80, 82-87 (except 861-2)//[b]

---

[a] A "//" separates Washington and U.S. SIC numbers where they do not perfectly match.

[b] U.S. 73 excludes 7396, 8921, 861, 862; includes 7692, 7694, pt. 7699
U.S. 72, 75-77 excludes 7692, 7694, pt. 7699, pt. 8099; includes 0722, 861, 862, 8921

Derivation of Location Quotient Estimates of Washington State Direct Intra-Regional Coefficients

Once the Washington and U.S. transactions tables were aggregated to similar 50-sector structures, the table of direct coefficients was computed for each by dividing each interindustry flow by the purchasing industry's gross output. Then, using the national direct coefficients table as a basis, two additional tables were derived, one using location quotients based on employment, and the other using location quotients based on output (Table 3-2). As previously described, for each industry i, $LQ_i = (N_i^R/N^R) / (N_i/N)$ is computed, where N is either employment or output, and R indicates the region. Each row of the U.S. table of direct coefficients is then multiplied by $\min [LQ_i, 1]$. The employment by sector was obtained primarily from the 1967 *County Business Patterns*. Gross output data by sector was taken from the respective transaction tables, whose sources are also standard government publications and readily available. These data sources are listed in the Appendix to this report, following the text.

At this point, then, we have four models of the Washington state economy:

(1) The survey-based "true" table for Washington, 1967 (WASH);

(2) The table based on location quotients using gross employment figures (EMPLQ), derived from the 1967 national table;

(3) The table based on location quotients using gross output figures (OUTPLQ), derived from the 1967 national table;

(4) The 50-sector national table whose sectors correspond to those of Washington, which can serve as what McMenamin and Haring call the "naive model" (NAIVE). Namely, we can use this model as a control, to see if the short-cut method of regionalizing national input-output tables provides any improvement over the naive use of simply the national model as an approximation of regional interindustry relationships.

A Slight Modification of the Application of Location Quotients Also Examined

At this point, it should be noted that the theory behind application of location quotients to national tables assumes that regional technology is approximately the same as that of the nation -- hence, if every sector in the region had enough capacity to supply any regional demand, no imports would be needed, and national and regional direct coefficients would be equal. If, however, sufficient production from within the region is not available, that portion of the interindustry requirements not met by

TABLE 3-2. EMPLOYMENT, OUTPUT, AND LOCATION QUOTIENTS FOR WASHINGTON AND THE U.S.

SECTOR	NAME	EMPL-WASH.	EMPL-U.S.	EMPL LQ	OUTPUT-WASH.[a]	OUTPUT-U.S.[a]	OUTPUT LQ
1	FIELD CROP	28950.	1918802.	0.94	210.3	21290.2	0.58
2	VEG&FRUIT	25280.	725213.	2.18	273.3	5914.7	2.71
3	LVSTK&PROD	16639.	1713604.	0.61	209.6	30637.7	0.52
4	FISH&FORES	16390.	880316.	1.16	242.7	5838.3	2.43
5	MEAT PRODS	3711.	305652.	0.76	294.0	22640.4	0.76
6	DAIRY PROD	3615.	228172.	0.99	208.7	13040.4	0.94
7	CANN&PRESR	7811.	219130.	2.23	394.2	9419.5	2.45
8	GRAIN MLLS	1453.	111373.	0.82	126.6	10053.5	0.74
9	BEVERAGES	2749.	218007.	0.79	239.0	12541.4	1.12
10	OTHR FOODS	5800.	515200.	0.70	179.8	23290.2	0.45
11	TEXTILES	645.	925159.	0.04	14.6	24808.2	0.03
12	APPAREL	5039.	1390846.	0.23	78.2	23832.3	0.19
13	MINING	2043.	594926.	0.21	59.1	24912.4	0.14
14	LOGGING	8507.	81990.	6.49	258.4	3049.4	4.96
15	SAWMILLS	17036.	194319.	5.48	419.3	4073.3	6.03
16	PLYWOOD	8066.	68507.	7.36	228.8	2000.1	6.70
17	OTHER WOOD	6983.	224319.	1.95	122.3	4750.1	1.51
18	FURN&FIXT	3349.	439118.	0.48	51.5	7878.7	0.38
19	PULPMILLS	3488.	15534.	14.04	185.6	1426.2	7.62
20	PAPER MLLS	8304.	140440.	3.70	292.0	6290.9	2.72
21	PAPBD MLLS	7600.	483900.	0.98	329.3	15909.8	1.21
22	PRINT&PUBS	10365.	1029091.	0.63	169.5	22118.0	0.45
23	INDUS CHEM	4845.	263932.	1.15	168.6	17041.0	0.58
24	OTHER CHEM	1952.	603269.	0.20	71.5	30088.9	0.14
25	PET REFINE	1329.	138512.	0.60	413.2	26974.8	0.90
26	GLAS&STONE	1000.	285000.	0.22	24.7	7185.1	0.20
27	CEMNT&CLAY	4551.	294952.	0.96	143.5	7662.3	1.09
28	IRON&STEEL	3178.	930450.	0.21	90.0	31723.4	0.17
29	NONFER MET	1524.	234700.	0.41	58.0	14774.0	0.23
30	ALUMINUM	8784.	133997.	4.10	556.1	6922.6	4.70
31	HEAVY METL	5266.	365665.	0.90	155.5	10614.7	0.86
32	LIGHT METL	2745.	905420.	0.19	106.4	26883.4	0.23

(continued)

Table 3-2. Employment, Output, and Location Quotients for Washington and the U.S. (Continued)

SECTOR	NAME	EMPL-WASH.	EMPL-U.S.	EMPL LQ	OUTPUT-WASH.[a]	OUTPUT-U.S.[a]	OUTPUT LQ
33	NONELC EQP	3624.	526145.	0.43	62.0	16603.6	0.22
34	MACH TOOL	2968.	564148.	0.33	96.3	12567.2	0.45
35	INDUS EQP	3860.	851878.	0.28	75.8	25178.2	0.18
36	ELEC MACH	5265.	1905171.	0.17	98.2	46759.2	0.12
37	AEROSPACE	95632.	826972.	7.23	2536.0	21993.1	6.75
38	MOTOR VEH	4230.	949302.	0.28	117.5	48381.6	0.14
39	SHIP BLDG	17500.	268697.	4.07	319.7	3135.5	5.97
40	OTHER MFG	6490.	2012791.	0.20	59.7	57632.7	0.06
41	TRANS SERV	60738.	3351400.	1.13	837.7	52825.3	0.93
42	UTILITIES	14688.	964008.	0.95	529.6	37348.0	0.83
43	COMMUNICAT	15504.	964155.	1.01	294.7	22511.5	0.77
44	CONSTRUCT	54945.	3004589.	1.14	2158.4	103280.0	1.22
45	ALL TRADE	213645.	13550368.	0.99	2516.3	163365.5	0.90
46	FINANCE	20577.	1369230.	0.94	353.0	23056.4	0.90
47	INSURANCE	15741.	1185209.	0.83	410.8	74593.1	0.32
48	REAL ESTAT	9846.	592170.	1.04	145.7	63555.9	0.13
49	BUS SERVIC	30849.	2061704.	0.94	460.0	56444.3	0.48
50	PERS SERV	117754.	7939657.	0.93	1189.1	93711.3	0.74
	OTHER	179034.	9439885.		20813.4	915396.4	
	TOTAL	1101907.	68906832.		39508.2	2313941.0	

[a] Output is X$$10^6$

regional producers must be met by imports; hence the direct regional coefficient between any two such industries would be decreased from the level of the national coefficient, and imports would become positive in the corresponding column of that matrix to supply the necessary additional flow. On the other hand, any production beyond regional (national technology) requirements is assumed to be exported, so that in such a case the direct regional interindustry coefficients are identical to the national coefficients, across that producing industry's row. From this interpretation it can be seen that regional coefficient estimates derived using location quotients can never exceed national coefficients. Clearly, this provides a built-in error when regional technology significantly varies from national relations, where regional coefficients do in fact exceed those of the national tables, a feature which cannot be captured by the standard location quotient method. To test whether a slight modification of the standard method could better estimate regional coefficients larger than national coefficients, an alternative method was used (STRICTLQ), where location quotients were multiplied across each row, regardless of whether they were greater than one. Thus, for those industries with location quotients greater than one, their regional coefficients were allowed to exceed the national coefficients. Two additional matrices of direct coefficients (one each for employment and output) were thus derived using this modification.

Estimation of Total Gross Output Using Estimated Tables

Next, the matrix of direct and indirect coefficients $(I-A)^{-1}$ was computed for each of the WASH, EMPLQ, OUTPLQ, and NAIVE models. Each of these matrices was then used to predict Washington gross outputs (already known) by using the familiar relationship.

$$X = (I-A)^{-1}F$$

where F = vector of final demands for regionally produced goods.

Simple Estimation with All Final Demand Given --

First, to test the predictive value of $(I-A)^{-1}$ with all final demand, including exports, given (known in fact from the survey table) the vector of gross outputs X was estimated by multiplying $(I-A)^{-1}$ by the given 1967 Washington final demands vector.

Simple Estimation with Inferred Exports as a Component of Final Demand --

To this point, we have considered only the estimation of the matrix of interindustry coefficients, and the use of a given final demands vector to calculate estimated total gross outputs. These are areas of primary concern. However, if we wish to construct other components of a regional input-output table, further computations may be performed.

Exports and imports may not be known or confidently projected in an actual practical application of location quotients, and they can be derived using only the known gross outputs vector and local final demand (and the estimated A-matrix). Schaffer and Chu (1969) describe a method of calculating exports and imports as residuals:

If $LQ_i > 1$, the industry is perceived as more important within the region than it is in the nation, so the sector can supply all intermediate demand and will export goods. For each such industry, exports

$$e_i = x_i - \sum_{j=1}^{s} x_{ij} - \sum_{f=1}^{t} y_{if},$$ where exports are computed as

the residual found by subtraction of all s intermediate industry demands and all t final demands for good i from gross output x_i of good i. No imports are needed to supplement interindustry transactions across that row.

The case where $LQ < 1$ is interpreted to mean that industry i's production is insufficient, and imports will be required to supplement interindustry transactions across that row. For any such flow, the imports required must be sufficient to compensate for the decrease in the direct coefficient from the national technology (assumed to approximate the regional technology), so for each element in such a row i,

$$m_{ij} = A_{ij}x_j - a_{ij}x_j$$

where m_{ij} = imports from industry i outside of the region to regional industry j
A_{ij} = national direct coefficient
a_{ij} = regional direct coefficient
x_j = regional total gross output.

By such a construction, row and column totals should hopefully equal the proper gross output, but this may not necessary occur, as exports will not necessarily be non-negative if $LQ_i \geq 1$, and will not necessarily be zero for $LQ_i < 1$. However, to adhere strictly to the theoretical method in this preliminary case, negative exports were allowed for $LQ_i \geq 1$, and zero exports were assumed for $LQ < 1$.

Clearly, such strict application of the simple location quotient technique will probably result in some cases of:

(1) ($LQ_i \geq 1$): rows summing correctly to total gross output but with negative exports;

(2) ($LQ_i < 1$): rows importing and therefore theoretically having zero exports not summing to total gross output.

If these exports are then incorporated in final demand, case (2) will cause $AX + F \neq X$, so $(I-A)^{-1}F \neq X$. So, for the two estimates of A derived using location quotients, we note that it is not likely that they will be able to accurately predict the total gross outputs vector X where exports as a component of final demand are imputed in such a strict interpretation. However, the results obtained are of interest in determining the usefulness of such strict direct application of the location quotient method. The balancing procedure described in the next section alters some of the strict assumptions of the location quotient method for the sake of internal consistency in a complete table, causing substantial modifications to the location quotient concept. (The NAIVE model, using national coefficients implicitly has $LQ_i = 1$ for all i, hence case (2) cannot occur and with exports (residuals) calculated for all rows, the predicted $X = (I-A)^{-1}F$ should be correct, although negative exports may still arise.)

Balancing the Tables

The inconsistencies of (1) and (2) above can be remedied according to either of two conventions. Each was applied to the original estimated tables.

Non-Negative Exports ($LQ \geq 1$) / Zero Exports ($LQ < 1$) --

The first alternative used to balance the estimated tables was to strictly follow the interpretation of $LQ_i < 1$ to imply insufficient production by industry i and hence none available for export. A balancing procedure, based on that of Schaffer and Chu, was used to modify the estimated direct coefficients matrix A and is described next.

Under the assumption that local final demand was to be taken as given, the preliminary exports were computed for all 50 rows as total output less local final demand less intermediate requirements based on the estimates of A. If $LQ_i \geq 1$ and exports were nonnegative, or if $LQ_i < 1$ and exports were zero, the row was left unaltered. If this was not the case, however, these residual-exports as a fraction of total intermediate output were distributed back into the A-matrix using a constant multiplier across the row. In this manner, if $LQ_i \geq 1$, negative exports were redistributed as imports as a constant fraction of row i, and exports were raised to zero to compensate for the decrease in the row of direct coefficients. Similarly, if $LQ_i < 1$ and preliminary residual exports were calculated as either positive or negative (i.e., nonzero), they were distributed back in a similar fashion over row i to yield zero exports and adjust row i's direct coefficients. If the regional coefficients rose to the level of the national (technical) coefficients, implying sufficient production to meet regional production requirements, no increase of coefficients above this level was allowed, and positive exports were then calculated.

It is to be noted that in fact such balancing is a form of the supply-demand pool method (see section VIII of this chapter). To summarize the preceding process, with local final demand taken as given and total gross output known,

1) if $LQ_i < 1$, exports must equal zero, hence all residual (total less final) intermediate sales per unit output must be distributed by some constant fraction over the estimated coefficients (which are already a constant fraction of national coefficients) unless or until the national coefficients are reached;

2) if $LQ_i \geq 1$, the national coefficients are used, and hence the residual exports (total output less known local final demand less known intermediate demand by national coefficients) must be the same as would be calculated with the NAIVE table as a basis. In the case where such exports would be calculated as negative, their redistribution over the intermediate row i coefficients would be simply an adjustment of the national coefficients.

From this it can be inferred that if the EMPLQ and OUTPLQ models both have $LQ_i \geq 1$ or both have $LQ_i < 1$ for a row i, their corresponding revised coefficients will be identical. Since it is likely that this situation actually holds for most rows, it is quite possible that the balanced EMPLQ and balanced OUTPLQ tables may be identical.

The new A-matrix was then used in $(I-A)^{-1}F = X$ to estimate total gross outputs using both total final demand known, and known local final demand plus imputed exports.

Non-Negative Exports ($LQ \geq 1$) / Positive Exports Allowed ($LQ < 1$) --
The second balancing method used was identical to that of the method above, except that (positive) exports were permitted even if $LQ_i < 1$. In actual fact, regional industries often do export products while they do not fully supply local intermediate demand. This slight modification permits such a characteristics to be captured.

A Comment on the Balancing Procedure --
It should be noted that this balancing procedure, necessary if one desires to construct a full consistent regional table, redefines the previous row multipliers (location quotients) of the national A-matrix so that the influence of location quotients is lost. By in fact taking total output and local final demand as given, and distributing the remainder of a row's production as in-ermediate sales, labelling any surplus above national technical requirements as exports, the initial application of location quotients becomes moot and NAIVE, EMPLQ, and OUTPLQ should all achieve the same results. (The allowance of positive exports with $LQ_i < 1$ in the second alternative does, however, allow variation.)

Summary of Procedures

Beginning with the application of the simple location quotient method to a national table of direct coefficients, estimates of a table of direct intraregional coefficients for the state of Washington were derived. Two sets of tables, based on employment and output location quotients respectively, were developed as was the NAIVE model of national coefficients. Strict application of $LQ_i > 1$ was also tested. For each of the original location quotient models and the NAIVE model, $(I-A)^{-1}$ was calculated, as were their estimates of total gross output using $X = (I-A)^{-1}F$, where given total final demand and then given local final demand plus inferred exports were used. The models were then balanced for consistency under two alternative conventions, and the revised A and $(I-A)^{-1}$ matrices were rederived.

For each of the steps, for all models, statistical comparisons were made between the estimated and actual figures. A description of statistical measures and the results of the tests are given in the following two sections.

V. MEASUREMENT OF ACCURACY OF ESTIMATES

Two types of analyses of the results are desired. First, in absolute terms, we would like to know how closely our estimates compare to the values of the "true" table. Secondly, we wish to note such comparisons in relative terms -- is one method better than another? by how much? are the additional data and computational requirements worth the effort?

Two types of estimates were produced by the previously discussed procedures: matrices of coefficients, and projections of flows. Each type can be analyzed using different tests. In all cases, comparison is made between the estimate and the "true" (survey-based) value.

Matrix Comparisons

The comparison of the elements of two matrices can be performed using a variety of tests (see in particular Morrison and Smith, 1974). Several measures were used in this analysis, but all have faults, which will be noted. Comparisons are made element-by-element over the entire matrix, for each row, and for each column. Computations over all elements in the matrix provide an overall view of the accuracy of the estimate, while statistics for each row and each column allow examination and evaluation by sector. In the descriptions below,

a_{ij} = the actual survey-based element

$\hat{a}_{ij}$ = the estimate of a_{ij}

n = number of elements considered:
2500 for whole-matrix calculations;
50 for row or column calculations

Summations Σ are taken over i (rows), j (columns), or i and j (all elements of the matrices) in all cases.

The statistical measures of comparison of the matrices of coefficients are:

(1) Mean Percentage Error (MPE) and its Standard Deviation (SDMPE)

$$\text{MPE} = \frac{\Sigma \dfrac{(|\hat{a}_{ij} - a_{ij}|)}{a_{ij}}}{n} \times 100\%$$

$$\text{SDMPE} = \sqrt{\frac{\Sigma|a_{ij} - \hat{a}_{ij}|^2}{n} - \left(\frac{\Sigma|a_{ij} - \hat{a}_{ij}|}{n}\right)^2} \times 100\%$$

Several observations can be made:

(a) The statistic is weighted inversely by the size of the true coefficient, so that an error in a small coefficient has a much larger effect on the MPE than would an error of the same size in a larger coefficient.

(b) If the actual $a_{ij} = 0$ while the predicted $\hat{a}_{ij} > 0$, then the fraction is undefined. In such cases, the (a_{ij}, $\hat{a}_{ij}$) pair is skipped in the computations.

Clearly, a small MPE indicates a closeness between the actual and predicted matrices.

The following five tests used were performed by Morrison and Smith (1974) to compare several regionalization methods in relative terms.

(2) Mean Absolute Difference (MAD)

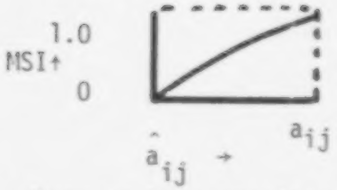
$$\text{MAD} = \frac{\Sigma|a_{ij} - \hat{a}_{ij}|}{n}$$

This expression is "useful to gain some impression of the magnitude of the deviations involved." Conclusions based on its results in absolute terms may be difficult to achieve, but its magnitude for one estimate as compared to that for another may be useful in determining relative similiarity -- which is "closer" to the "true" table. A small MAD indicates closeness to the "true" table.

(3) Mean Similarity Index (MSI)

$$\text{MSI} = \frac{\Sigma\left(1 - \dfrac{|a_{ij} - \hat{a}_{ij}|}{a_{ij} + \hat{a}_{ij}}\right)}{n}$$

Note that $0 \leq \text{MSI} \leq 1$, with 1 indicating perfect similarity and 0 indicating no similarity. For a given a_{ij}, the relationship between $\hat{a}_{ij}$ and MSI can be graphed as



Clearly an MSI as close to 1 as possible is desired.

(4) Correlation Coefficient (R) (also R^2)

$$R^2 = \frac{\left(\Sigma a_{ij}\hat{a}_{ij} - \frac{\Sigma a_{ij}\,\Sigma \hat{a}_{ij}}{n}\right)^2}{\left(\Sigma a_{ij}^{\;2} - \frac{(\Sigma a_{ij})^2}{n}\right)\left(\Sigma \hat{a}_{ij}^{\;2} - \frac{(\Sigma \hat{a}_{ij})^2}{n}\right)}$$

R also ranges from 0 to 1, with 1 indicating a perfect correlation between actual and estimated coefficients. However, R^2 may equal one even if the correlation between the two matrices is negative. Thus high R^2 values do not guarantee that the two matrices under comparison are necessarily identical, only that a straight-line function can approximate the relationship between them.

(5) Information Content (INFO)

This statistic is borrowed from information theory. The estimated table is considered a forecast of the actual table, and the additional information which is gained upon receiving the second (actual) table is expressed as

$$I(A:\hat{A}) = \Sigma \left(a_{ij} \left| \log_2 \frac{a_{ij}}{\hat{a}_{ij}} \right|\right)$$

So, if the estimate is accurate and closely mirrors the actual table, then the additional information received from the true table should be small. Hence, a small value of INFO means a close approximation of the true table by the estimate. Several observations are in order:

(a) The value of each term in INFO is weighted by the actual coefficient a_{ij}, placing more importance on coefficients of larger size.

(b) An absolute interpretation of its value is difficult, so that it is primarily useful for making relative comparisons between methods.

(c) In a situation converse to that of MPE, INFO is undefined if the estimated value $\hat{a}_{ij}$ is zero while the actual coefficient a_{ij} is positive. Because of this, such elements a_{ij} cannot be included in computation of INFO.

(d) For the case where $a_{ij} = 0$ and $\hat{a}_{ij} > 0$, the term in INFO equals zero (a characteristic which can be found by applying l'Hôpital's rule). That is, even if an estimated coefficient is non-zero while the actual element is zero, no additional information is contained in the actual coefficient. Clearly, when such an error is not included in the statistic, INFO becomes even weaker as a measure of comparison between two matrices.

(6) "Chi-Square" - x^2 (CHISQ)

$$\text{CHISQ} = \Sigma \frac{(a_{ij} - \hat{a}_{ij})^2}{a_{ij}}$$

Note that this statistic also is weighted inversely by the size of the true coefficient. Also, it is undefined (infinitely large) when $a_{ij} = 0$ and $\hat{a}_{ij} > 0$. Schaffer and Chu and McMenamin and Haring, computing over columns, treat such cases by combining them into one cell to form a new technical coefficient, and redefining the actual flows in the zero cells as $49,000, the largest number which would be registered as zero in the survey-based table. This admittedly arbitrary "adjustment" is credited with avoiding zero division and extreme chi-square values. However, extreme chi-square values can occur almost as easily with very small actually-observed coefficients. Such a procedure of arbitrarily establishing a minute flow in order to include a chi-square term appears to be more of a testimony to the faults of the chi-square statistic than to be a consistent method of overriding those flaws. For this reason, such coefficient pairs were not included in the chi-square calculations; however, their number and average magnitude were recorded, to provide some measures of those coefficients not included in the calculations. Morrison and Smith also chose not to include them in their chi-square calculations.

Even if Schaeffer and Chu's method for counting zero flows were used, however, it should be noted that this statistic which the previous studies call chi-square does not in fact have a chi-square distribution, unless the a_{ij}, $\hat{a}_{ij}$ which were expressed in terms of percentages by McMenamin and Haring can in some way be interpreted in the customary manner as frequencies rather than proportions. To arbitrarily use percentages rather than fractions (by multiplication by 100) multiplies CHISQ by 100, and the results may seem "reasonable" by evaluating levels of significance. However, such multiplication by 100 implies that a sample of 100 points (e.g., dollars) of inputs were purchased by industry j and the number of dollars sold from each industry i to j was found to be that particular amount 100 x a_{ij}. However, an analogous interpretation could use any units at all to multiply a_{ij} and $\hat{a}_{ij}$. (e.g., millions of dollars input per millions of dollars output) and implied sample size would increase as would the CHISQ statistic, by that multiplier. Hence, while CHISQ clearly measures a sum of weighted squared deviations which is useful in comparative terms, the interpretation of a significance level read off of a chi-square table appears to have little meaning, as the decision to force the coefficients (fractions) to appear as frequencies introduces an arbitrary factor into the "chi-square statistic" with slim theoretical justification. Calculation of these "pseudo-significance levels" may only be useful in comparing Schaffer and Chu's and McMenamin and Haring's number of "acceptable" sectors at "α" = .05, with those found by using the same tests on our derived matrices.

Flow Vectors Comparisons

Comparisons of the vectors of flows predicted by the various estimated tables can be made in much the same way as was done with the matrices. Mean Absolute Difference, Mean Percentage Error, and Mean Similarity Index can be used as previously defined, with no loss in their intuitive interpretations. The Information Content can still be used for relative comparisons between methods. While not distributed with chi-square distribution, the formula

$$\text{"CHISQ"} = \Sigma \frac{(z_j - \hat{z}_j)^2}{z_j}$$

for any vector of flows z can also provide an intuitive impression of the magnitude of deviations in values derived by different methods as a "weighted sum of squared deviations" term. The linear least squares fit is probably the set of statistics of most importance and calculations are made for R, R^2, standard error of estimate (STDERR), slope (BHAT), and y-intercept (AHAT) of the least squares regression line. Ideally, BHAT $\approx$ 1, and AHAT $\approx$ 0.

VI. ANALYSIS OF RESULTS

As previously stated, our main concern in the test of the use of location quotients to regionalize national input-output tables was to accurately derive the matrix of direct coefficients which, given a vector of regional final demands, can be used to calculate total gross output by industry:

$$(I-A)^{-1} F = X.$$

By such a procedure, regional outputs can be predicted by a means other than allocating national production. The various models described in Section IV were analyzed with a variety of statistical measures (detailed in Section V). The coefficients of the A and $(I-A)^{-1}$ matrices and the predicted total gross outputs, exports, and imports vectors for each of the models were compared with those of the "true" survey-based Washington state table. The results are described below.

The Initial Estimates Using Location Quotients

Originally, it will be recalled, three estimates of the 1967 Washington state input-output table were derived: 1) EMPLQ, estimated by multiplication of the direct coefficients table of the national economy by location quotients based on employment; 2) OUTPLQ, estimated by multiplication of the national table by location quotients based on output; and 3) NAIVE, the national table which assumes the state economy is identical to the national economy.

Estimates of the Direct Coefficients Matrix --

The element-by-element comparison of direct coefficients from the various models indicated less than impressive results for the matrix as a whole. Each model's estimated matrix (NAIVE, EMPLQ, OUTPLQ) was compared to that of the true table (WASH) (see Table 3-3). In absolute terms, no measure yielded very impressive results (all $R^2 < .54$, all MSI < .48, all MPE > 135%). However, as hypothesized, improvement was noted of the EMPLQ model's estimated coefficients over the NAIVE (U.S.) model, and of the OUTPLQ's model over those of EMPLQ, in all tests except information content. Mean Absolute Difference was reduced by 45% and CHISQ by about 75% in using OUTPLQ rather than NAIVE. The anomaly of the Information Content's value is not very disturbing, as we have previously discussed its shortcomings.

Examination of the individual sectors by column indicated significant disparities between estimated and actual coefficients. Even for the closest estimate (OUTPLQ), no MPE was less than 23%, only 7 sectors had less than 50% MPE, and 22 had values of MPE greater than 100%. Only one column's mean similarity index was greater than 0.7 (on a scale of 0 to 1). Ordered by output, 15 of the highest 21 sectors had R > .8 for OUTPLQ, compared to 12 for EMPLQ, and 13 for NAIVE, but no other statistics provided similar encouragement.

Also, in order to compare these results with previous studies, Schaffer and Chu's "chi-square" test was used, using the "$49,000" convention for zero cells, and "pseudo-significance levels" determined from the CHISQ statistics calculated (with the coefficients expressed as percentages). Of the 50 sectors, only three were acceptable at "α" = .05 in the NAIVE model, 16 in the EMPLQ model, and 25 in the OUTPLQ model. It is recalled that, using this test, Schaffer and Chu obtained only 6 of 23 sectors significant at "α" = .05 using location quotients, while McMenamin and Haring, in projecting from the 1963 Washington table to that of 1967, obtained only 10 such "acceptable" sectors in the NAIVE model, and 12

TABLE 3-3. COMPARISON BETWEEN DIRECT COEFFICIENTS TABLES OBTAINED FROM SURVEY AND SHORT-CUT METHODS

Model Compared with True Washington Table

Statistic	NAIVE (National)	EMPLQ	(with strict LQ)	OUTPLQ	(with strict LQ)
R	.6485	.7169	.4407	.7328	.4858
R^2	.4205	.5140	.1943	.5370	.2360
MAD	.007734	.005197	.010863	.004287	.009413
MSI	.4670	.4679	.4685	.4753	.4664
MPE [a] (s.d [a])	314.16 (1196.59)	161.91 (618.20)	228.46 (1071.64)	135.02 (552.74)	205.56 (1028.36)
CHISQ [a] (zero) [a] (avg) [a]	255.049 (865) (.00555)	77.176 (864) (.003001)	640.432 (864) (.004693)	62.976 (864) (.001883)	590.445 (864) (.003468)
INFO [a,b] (zero) [b] (avg) [b]	11.422 (41) (.004795)	12.708 (41) (.004795)	15.259 (41) (.004795)	13.191 (41) (.004795)	15.970 (41) (.004795)

[a] When $a_{ij}=0$ and $\hat{a}_{ij}>0$, MPE and CHISQ are undefined, and INFO is zero, for that element. zero=number of such elements and avg=average of such elements.

[b] When $a_{ij}>0$ and $\hat{a}_{ij}=0$, INFO is undefined, for that element. zero=number of such elements and avg=average of such elements.

using RAS and H-M, out of 27 sectors. Comparing this to the fully one-half of our 50 sectors estimated as "acceptable," this suggests a substantial improvement in our estimation of the Washington economy using the national table.

Hence, our results indicated that the coefficients displayed a definite increase in accuracy when location quotients were applied to the national table, particularly when output rather than employment was used. For example, from NAIVE to EMPLQ to OUTPLQ, the mean percentage error decreased from 314% to 162% to 135%. However, these raw figures taken by themselves still do not indicate very close estimation. In fact the correlation coefficient ranged, respectively, from .649 to .717 to .733.

For the test of whether strict application of location quotients (multiplying across rows by LQ_1 regardless of $LQ_1 > 1$), statistics indicated poorer estimates of the coefficients resulted ("strict" column in Table 3-3).

Estimates of Direct and Indirect Coefficients --

The practical test of the estimated A-matrix is its predictive value, obtaining gross outputs using A and the final demands vector: $X = (I-A)^{-1}F$, where $(I-A)^{-1}$ is the matrix of direct and indirect coefficients.

The A-matrices found by WASH, NAIVE, EMPLQ, and OUTPLQ were used to derive the tables of direct and indirect coefficients, $(I-A)^{-1}$ (see Table 3-4).

Again significant improvement was evidenced in progressing from NAIVE to EMPLQ to OUTPLQ. For example, mean absolute difference, reduced by a factor of about two-thirds in moving from NAIVE to OUTPLQ, measured .0056, on the same order as its value for the A-matrix (.0043). However, the very high correlation coefficients -- .976, .991, and .993 -- are likely due to the extremes of coefficient values, concentrating either at small fractions or greater than one on the matrix's diagonal, allowing a least squares regression to explain much of the variation about the mean. Also, even the model yielding the best results, OUTPLQ, gives only MSI = .4924, a statistic which considers each pair of elements separately and so should not be exaggerated by the extreme elements along the diagonals. The results of comparing the known values to those derived by the various methods of estimation again indicate improvement with progression to the output-location quotient model. However, absolute differences appear to be too high to confidently accept the models' ability to depict accurately each of the regional economy's interindustry relationships.

Estimates of Total Gross Outputs --

Next, the gross outputs were predicted by each estimate of the A matrix using the assumption of all final demand (including exports) as given. In Table 3-5, results are very impressive in the prediction of the total gross outputs vector. Dramatic improvement was seen in movement from the NAIVE to the EMPLQ to the OUTPLQ, but here, unlike the case of the coefficient matrix comparisons, the measures taken in absolute as well as

TABLE 3-4. COMPARISON BETWEEN $(I-A)^{-1}$ MATRICES OBTAINED FROM SURVEY AND LOCATION QUOTIENTS METHODS

Model Compared with True Washington Table

Statistic	NAIVE (National)	EMPLQ	OUTPLQ
R	.9760	.9908	.9928
R^2	.9525	.9818	.9857
MAD	.015979	.007657	.005624
MSI	.3004	.4388	.4924
MPE[a]	13394.86	4212.93	2587.67
(s.d.)	(99989.25)	(52577.44)	(31390.94)
CHISQ [a]	12832.965	2388.330	987.504
(zero) [a]	(2)	(2)	(2)
(avg)	(.250980)	(.000354)	(.000145)
INFO [a,b]	21.615	17.050	16.721
(zero) [b]	(0)	(0)	(0)
(avg) [b]	---	---	---

[a] When $a_{ij}=0$ and $\hat{a}_{ij}>0$, MPE and CHISQ are undefined, and INFO is zero, for that element. zero=number of such elements and avg=average of such elements.

[b] When $a_{ij}>0$ and $\hat{a}_{ij}=0$, INFO is undefined, for that element. zero=number of such elements and avg=average of such elements.

TABLE 3-5. COMPARISON BETWEEN ACTUAL AND ESTIMATED OUTPUT COMPUTED WITH FINAL DEMAND AS GIVEN

Model Compared with True Washington Table

Statistic	NAIVE (National)	EMPLQ	OUTPLQ
Mean Percentage Error	161.95 (σ=239.28)	32.79 (σ=53.59)	16.63 (σ=17.17)
Mean Similarity Index	.7053	.8842	.9256
Linear Least Squares Fit			
R	.9651	.9883	.9941
R^2	.9314	.9768	.9882
Std. Error of Est.[a]	171.2415	100.6733	70.4190
Estimated Slope[a]	1.12617	1.16428	1.15134
Estimated y-int.[a]	148.865	21.544	-11.527
Mean Absolute Difference[a]	197.2	85.6	54.9
Information Content[a]	8752.324	5008.187	3593.715
Sum of Weighted Squared Differences ("CHISQ") [a]	26022.262	3173.135	727.466
Gross Output for the 50 Sectors of Washington [a]			
Actual Output=18694.8[a]	28496.6	22843.0	20947.7

[a] Output is in $\$10^6$

comparative terms indicate good estimates. Mean percentage error (and its standard deviation) drops from 161.95% (239.28%) to 32.79% (53.59%) to 16.63% (17.17%), from the NAIVE to OUTPLQ estimates, an improvement of almost 90%. The mean similarity index of OUTPLQ's estimates reaches .9256 on a scale of zero to one. Probably the most significant result is the fit of the regression line for the estimated output as a function of the actual. Ideally, we wish $R^2 = 1$, slope of line = 1, and y-intercept = 0, to indicate perfect estimation. The NAIVE model predicts gross outputs with $R^2 = .9314$, but its estimated y-intercept is 149 and slope is 1.126, resulting in a consistent overestimate of values, with a standard error of 171. An error of 41% is present between actual and predicted total gross output. The EMPLQ dramatically improves R^2 to .9768, standard error of estimate of 101, and estimated y-intercept of 21.5. The slope again approaches 1, and total output for the 50-sector is overstated by only 22%. All other statistics show great improvement. Finally, the OUTPLQ model results in a mean percentage error of only 16.63%. A value of $R^2 = .9882$ is achieved by a line of y-intercept -11.5 and slope 1.15. The estimated total output is only 12% from the actual value for the 50-sectors, and all other measures are similarly dramatically encouraging.[6] These results indicate that, despite some inaccuracy in the coefficients matrices estimated with location quotients, these estimates can provide quite accurate estimates of total gross outputs X.

Estimates of Other Portions of the Table --

While the derivation of the A and $(I-A)^{-1}$ matrices and the estimated total gross outputs vector which follows directly were the primary goal of our study, other portions of the regional input-output table can be derived as discussed in Section IV of this chapter. The ability of the A-matrix to estimate the vector of intermediate outputs was tested, and exports and imports by sector were derived.

Intermediate outputs were calculated as AX, where A was the estimated matrix and X was the true vector of gross outputs. The results in Table 3-6 indicate continuous improvement as we progress from NAIVE to EMPLQ to the OUTPLQ model, except for a small dip in R^2. In fact, the improvement

---

[6] Because all three estimates overstate total output while a regression line indicates high correlation between actual and estimated values, the possibility of some type of bias in the procedure was raised. However, no evidence to explain such a bias could be determined. The NAIVE model clearly must overstate predicted output because all national coefficients must be greater than or equal to those derived using location quotients. The location quotient adjustments (reductions) of the national coefficients can be expected to reduce predicted gross output. No direct or implicit causes of bias could be determined in the estimation procedure, and in fact, the evidence provided by only the two cases is probably not a valid one upon which to conclude the existence of bias. Further, the close fit of a least squares regression line coupled with a slight overstatement of estimated values suggests that a possible adjustment factor could be easily used to adjust other estimates derived from the model with little additional computation if such cases were to occur consistently.

TABLE 3-6. COMPARISON BETWEEN INTERMEDIATE OUTPUTS OBTAINED FROM SURVEY AND SHORT-CUT METHODS

	Model Compared with True Washington Table		
Statistic	NAIVE (National)	EMPLQ	OUTPLQ
Mean Percentage Error	563.73 (σ=1065.11)	222.80 (σ=447.93)	156.14 (σ=305.22)
Mean Similarity Index	.5592	.6981	.7289
Linear Least Squares Fit			
R	.6502	.7128	.6864
R^2	.4227	.5081	.4711
Std. Error of Est.[a]	109.8868	101.4799	87.9514
Estimated Slope[a]	1.18282	1.29725	1.04408
Estimated y-int.[a]	84.71695	30.01765	26.58334
Mean Absolute Difference[a]	101.3	59.5	44.7
Information Content[a]	3383.754	2383.192	2112.780
Sum of Weighted Squared Differences ("CHISQ")[a]	35950.516	10653.629	6517.984
Gross Intermediate Output for the 50 Sectors of Washington[a]			
Actual Output = 3952.2[a]	8910.6	6627.9	5455.6

[a] Output is in $\$10^6$.

is quite substantial, with Mean Percentage Error decreasing by over a factor of two-thirds. However, even the value of 156.14% MPE for OUTPLQ must be considered too high to be acceptable. Mean Absolute Difference decreases by over one-half, and the overestimate of total intermediate output decreases by almost 70%, from 4958.4 to 1503.4 x $\$10^6$ in moving from the NAIVE to the OUTPLQ model. But again, even this OUTPLQ estimate overstates actual intermediate output by almost 40%, a sizable error. This set of statistics thus confirms our previous findings on the A-matrix coefficients in showing a substantial degree of discrepancy in interindustry relationships. The good predictive performance of our estimates of $(I-A)^{-1}$ in estimating total outputs, however, remains as an encouraging sign for our main goal of estimating regional production by modifying national input-output tables.

As described previously, while $(I-A)^{-1}$ can be used to determine X, given final demands F, and AX can be used to find intermediate outputs, exports and imports can also be found as residuals using the precise interpretation of location quotients. (Exports are total less intermediate output less local final demand, if $LQ_i \geq 1$, and imports are zero; or exports are zero and imports are national technology requirements not met by local coefficients in the case of $LQ_i < 1$.)

As can be seen in Table 3-7, exports calculated as residuals in all three models are highly correlated (R^2 = .8537, .9146, .9294 for NAIVE, EMPLQ, and OUTPLQ, respectively), but they greatly underestimate the true values.

NAIVE gives $R^2 = 0$ for imports, because its implicit $LQ_i = 1$ presumes that zero imports are required (see Table 3-7). However, once again EMPLQ (R^2 = .9308) and OUTPLQ (.8364) are highly correlated with the true Washington values. They too are understated compared to the true values.

Replacing the given Washington exports component of final demand with those calculated above, and recalculating $X = (I-A)^{-1}F$, the NAIVE model, as expected, now perfectly estimates X, because the calculation of exports for each row (since $LQ_i = 1$, $i = 1, \ldots, 50$) as a residual balances the table so $X = (I-A)^{-1}F$ is true by definition. The EMPLQ and OUTPLQ estimates now underestimate X (by approximately the same amount previously overestimated), as all exports are set to zero in rows with $LQ_i < 1$.

Modifications: Balancing the Tables

Balanced Tables with Zero Exports in Rows with $LQ_i < 1$ --
The balancing procedure, described earlier, was performed to achieve consistency of results by removing negative exports and balancing where rows of $LQ_i < 1$. As explained previously in Section IV, this procedure, by either

1) computing exports where all intraregional interindustry requirements (national technology) are met, as are the given final demand and total output; or

TABLE 3-7. COMPARISON BETWEEN EXPORTS AND IMPORTS OBTAINED FROM SURVEY AND SHORT-CUT METHODS

Model Compared with True Washington Table

	EXPORTS			IMPORTS		
Statistic	NAIVE	EMPLQ	OUTPLQ	NAIVE[c]	EMPLQ	OUTPLQ
MPE (s.d.)	1198.22 (2981.84)	619.01 (2594.09)	283.46 (1399.01)	100.0 (.35)	43.78 (28.71)	56.20 (68.99)
MSI	b	b	b	0.0	.7186	.7739
Linear Least Squares Fit						
R	.9240	.9563	.9641	0.0	.9648	.9146
R^2	.8537	.9146	.9294	0.0	.9308	.8364
Std. Error of Est. [a]	98.4109	68.1977	59.0128	--	14.9351	40.4107
Est. Slope [a]	.82336	.77292	.74156	--	.35557	.44441
Est. y-Intercept [a]	-77.94252	-35.95230	-25.04193	--	11.61668	26.55054
MAD [a]	101.3	65.3	58.2	95.7	52.5	45.3
INFO [a]	3453.575	1790.711	1913.236	--	6020.336	4238.055
Sum of Weighted Squared Differences ("CHISQ") [a]	124314.187	73482.812	12255.336	--	1626.884	1514.051
Predicted Gross Total [a]	1052.1	2848.4	3205.4	0.0	2283.0	3455.0
(Actual Exports=6011.0) (Actual Imports=4787.2)						

[a] Output is in $\$10^6$.

[b] Possible negative values for exports cannot be included in MSI calculations.

[c] All predicted imports = 0.

2) with given final demand and total gross output met, distributing the remainder of production across the row, with exports as zero,

must result in the EMPLQ, OUTPLQ, and NAIVE models becoming identical. As a consequence, the results in Table 3-8 hold for all three models in comparison to the true WASH table. The revised A and $(I-A)^{-1}$ matrices can be seen to be a clear improvement over the originally-derived NAIVE model. Some statistics improve, some worsen compared to the original EMPLQ, but the balanced tables are still inferior in most tests to the accuracy of the original OUTPLQ. Exports were better estimated than before balancing where negative exports were permitted. Because of the balancing, X must again be perfectly estimated using the revised A when final demands include inferred exports. With all final demands (including exports) used as given in the survey-based table, estimated output, although overstated, has very high correlation with the actual values.

Again, however, it is to be remembered that the adjustments used in balancing nullify modifications made by the location quotients and distort the initially derived estimated A-matrices.

Balanced Tables Allowing Positive Exports in Rows with $LQ_1 < 1$ --

As explained in section IV, the originally-derived tables were also balanced in a similar way except that positive exports were allowed for rows with $LQ_i < 1$. As a result, the NAIVE model must be unchanged. The minor alterations to EMPLQ and OUTPLQ did provide a small measure of improvement over the other balancing procedure, and a slight improvement over the original tables, but at the expense of the added manipulations to recreate a full given table rather than the simple location quotient estimation of the input-output matrix.

Summary of Results of Balancing --

As we have tried to emphasize, our main goal was to estimate the A-matrix (and therefore $(I-A)^{-1}$ and total gross outputs X, given final demands F). The balancing procedure, while an interesting modification, nullifies the effect of location quotients in its attempt to estimate a complete consistent table, and complicates our simple short-cut method. In the estimation of the A and $(I-A)^{-1}$ matrices, the improvement due to the use of the balancing procedure was rather small compared to the initial estimates based on location quotients. However, it does permit estimates of other portions of a regional table, in cases where such information is desired.

VII. CONCLUSIONS AND SUMMARY

In this description of our effort to examine the ability of two types of location quotients to adjust national input-output tables to reflect intraregional interindustry transactions, we have attempted to highlight our efforts at improving the methods used in previous studies.

TABLE 3-8. COMPARISON BETWEEN TABLES OBTAINED FROM SURVEY AND BALANCING PROCEDURE[a]

Portion of Table[b]

Statistic	A	$(I-A)^{-1}$	$X=(I-A)^{-1}F$ [c]	Intermediate Outputs Vector	Exports Vector	Imports Vector
Linear Least Squares Fit						
R	.7183	.9912	.9925	.7308	.9867	.9152
R^2	.5159	.9826	.9850	.5340	.9736	.8375
Std. Err. of Est.			77.4086	78.5890	36.0150	35.7976
Slope			1.11973	1.05824	.75762	.39525
y-int.			20.4040	37.75821	-13.32056	18.87143
MSI	.4895	.4593	.8835	.7007	.3950	.7449
MAD	.004851	.006824	66.4	44.4	44.5	49.0
MPE (s.d.)	154.40 (625.10)	4973.41 (59873.09)	29.18 (27.39)	266.07 (681.55)	66.07 (37.21)	49.95 (46.13)
Sum of Wtd. Sqd. Diff. ("CHISQ")	78.840	2622.307	1195.043	8723.844	1297.396	1509.168
INFO	12.164	16.762	4142.445	1772.479	3453.575	5179.305
Gross Total	---	---	21953.6	6070.3	3888.0	2835.7
(actual from survey table)	---	---	(18694.8)	(3952.2)	(6011.0)	(4787.2)

[a] NAIVE, EMPLQ, and OUTPLQ all yield the same results.

[b] Output is in $\$10^6$.

[c] F is given Washington final demands vector.

First, the input-output matrices estimated using location quotients were shown to be a significant improvement over the naive assumption of national interindustry relationships within the region.

Second, location quotients based on output were shown to provide better adjustment factors for the national table than did those in which employment was the measure of importance of an industry.

Third, the application of a large number of previously-used tests of the accuracy of the estimated table allowed some degree of comparison to be made with previous tests. In this way, for example, the effects of a large number of sectors (and hence more homogeneous industry) and the use of output as opposed to employment in computation of location quotients could be examined. It was found in virtually all tests that output data provided better results than employment. And, while in general the results of the present study indicated improved estimation ability over previous studies, contrary conclusions based on some comparisons or other tests may be explained by the increased cumulative error in the summary comparative statistics due simply to the increase in data points (number of sectors, squared).

Finally, in the desire to estimate accurately regional outputs using a short-cut method to regionalize national input-output tables, the most important goal is to be able to accurately estimate regional outputs. Our application of location quotients achieved quite good results using the given Washington final demand vector.

By removing some handicaps suffered by other studies, and achieving improved results, it is tempting to suggest that further improvements may be attained if the problems we could not remove were eventually eliminated. Specifically, definitional differences in sectors and conventions between Washington and national tables may account for additional error. The transfers of secondary products and imports within the national table in particular could be expected to influence the effect of location quotient adjustments. Further, the consistent overstatement of all three estimates suggest some possible bias. It may be feasible to apply some sort of adjustment factors to account for this bias. While these problems have not yet been solved, the success achieved in the present study suggests that location quotients can indeed be viable means of regionalizing national input-output tables.

When an individual state or region wants to explore the effects of national economic trends, environmental standards or energy development on the economy and environment of its state or region, the preparation of a state or regional I-O table by the location quotient method may be warranted. While the resulting I-O table may not accurately reflect a survey-based regional table, it may be sufficiently close to permit approximate estimates of the direct and indirect effects of exogeneous developments on the regional economy. Thus the location quotient method may provide an additional tool that allows a first order estimate of effects that an elaborate interindustry, interregional model (along with current and projected data bases) could provide if it were available. In other words, location quotients would be, in our view, an acceptable method for regional analysis from the perspective of a regional or state administration, until such time as a satisfactory regional model becomes available.

VIII. APPENDIX: A REVIEW OF OTHER METHODS OF REGIONALIZING NATIONAL INPUT-OUTPUT TABLES

Short-cut methods other than the use of location quotients have been employed to derive estimates of regional input-output tables from national tables. A survey of those methods is presented below.

Commodity-Balance Method or Supply-Demand Pool Technique (Morrison and Smith, 1974)

The commodity balance method, or "Supply-Demand Pool Technique" (SDP) uses the balance $b_i = x_i^R - d_i$, where x_i^R = local output of good i, and d_i = local requirements of good i ($d_i = \Sigma_j x_j^R a_{ij}$). If $b_i < 0$, then local supply covers local demand, and the national technical coefficients are used in row i of the regional trade coefficients matrix. If $b_i > 0$, then the national technical coefficients are reduced by a factor of x_i^R / d_i. Finally, using these results, imports or exports are computed.

RAS Iterative Technique

The RAS iterative technique (RAS) was developed to project future input-output tables from a base period matrix, but can be applied to project regional tables from national tables. Survey data is required to obtain intermediate input and output data by sector. Regional intermediate output, intermediate input, and gross output vectors, u, v, and x constrain input-output coefficients-matrix A iteratively in the following manner, where i = the identity vector and a circumflex indicates a diagonalized vector:

$u_1 = A_1 x$, to first approximate regional intermediate output with national coefficients;

$A_2 = \hat{u}\, \hat{u}_1^{-1} A_1$, to satisfy row constraints;

$v_1 = i\, \hat{x}\, A_2^T$, to estimate intermediate inputs;

$A_3 = A_2\, \hat{v}\, \hat{v}_1^{-1}$, to satisfy column constraints;

and the new estimate of A is used to repeat the process until convergence is reached and both sets of constraints are fulfilled (Morrison and Smith, 1974).

Czamanski and Malizia --

Czamanski and Malizia (1969) used the RAS method in this manner to derive a regional input-output table for Washington state (1963) for which a survey-based table was available. In order to test which sectors could be satisfactorily estimated by RAS and which require surveys, the authors tested six cases of various sector aggregations and relative price adjustments. Convergence in the computations was so rapid that the authors concluded that no iterations need be made after the first one.

Using as measures of deviation between the estimated and survey-based models standard deviation and mean percentage error of the coefficients and information theory "I" values[7], the results were examined. In all six cases, there were large deviations between real and estimated values. Reducing the number of sectors from 43 to 36 by grouping the tertiary sectors yielded some improvement, as Czamanski and Malizia had hypothesized that major errors could have resulted within them due to differences in routing practices. The wide range of activities in the primary sectors was also expected to cause poor results. In fact, the errors in these sectors were more than two standard deviations larger than the mean errors of the entire matrix. High regional specialization was the third factor predicted to cause errors. In fact, six of the 36 sectors had location quotients (using employment) of 2.0 or more. Regional specialization was observed to appear to imply a different technology than the national average, with an exception of the aerospace industry of which Washington has a large proportion of the national total.

The measures of error, mean percentage error (and its standard deviation), ranged from 58.65 (and 2.211) to 80.81 (and 6.314), while "I" values were 6.279 to 54.169. Only when the 36-sector matrix was reduced by four primary industries and four regionally specialized industries to 28 sectors, did the "I" value reduce to 0.779, which the authors claim is "certainly an acceptable level by any standard" (Czamanski and Malizia, 1969, p. 73). From this they conclude that with considerable adjustments of national input-output tables, "acceptable results can be achieved by the methods tried on the Washington state table." However, the mean percentage error was 38.93 (with standard deviation 2.160). Miernyk, in commenting on this paper, notes that the error is "uncomfortably high" (Miernyk, 1969, p. 82). He also points out that in steps performed to reach this level of accuracy, six sectors were dropped (which were equivalent to twelve in the Washington table), and ten sectors were aggregated into two, transforming a 54 sector Washington table into 34 sectors. This means that by Czamanski and Malizia's proposals, field surveys would be required for 20 sectors

---

[7] Where a_{ij} = true coefficient and a^*_{ij} = estimated coefficient,

$$I(A : A^*) = \sum_i \sum_j | a_{ij} \log_2 (a_{ij} / a^*_{ij})| ;$$

this expression is undefined when $a^*_{ij} = 0$ and $a_{ij} \neq 0$, a problem which must be arbitrarily resolved in computations.

(accounting for 13% of intermediate sales, 12% of final demand, and 28% of gross output). Further, Miernyk points out that the sectors in the combined coefficients are important analytically, and that if surveys are needed for 28% of output, the other key sectors should be identified and surveyed as well. Miernyk's conclusion is that "Czamanski and Malizia performed a signal service by measuring the large errors that can result from short-cut methods. They have provided the empirical evidence which has been missing up to now to support a priori arguments that adjusted national coefficients do not reflect true structural differences among regions. (Miernyk, 1960, p. 82).

McMenamin and Haring --

McMenamin and Haring (1974) describe other applications of the RAS method. The Belgian national economy was projected from 1953 to 1959 by Paelinck and Waelbroeck. Bates and Bacharach explained errors from three causes. Sectors were highly aggregated, a fault of the study and not RAS. (It is interesting to note that Czamanski and Malizia were forced to aggregate some sectors to improve their results.) A second cause, attributable to RAS, is the effect of uniform substitution assumed over all sectors upon a change in the requirements of one. The third problem with RAS is the ripple effect -- an error in one estimate generates errors in the rest of the table. Paelinck and Waelbroeck obtained a more accurate table by removing cells seen to be troublesome in advance, recomputing with RAS, and inserting the survey values for the 1959 table. McMenamin and Haring point out several facts about the Belgian test of RAS. Few changes occurred in the coefficients between 1953 and 1959. Of the 270 nonzero coefficients, 238 had an absolute difference of less than .005 between the estimated and survey tables. The actual values for total gross output, total intermediate input, and total intermediate output from the 1959 table were used in the iterations, avoiding testing the effect of the alternative use of independent estimates. Finally, some cells were replaced by 1959 survey values rather than independent estimates. Thus, the test represents an upper limit on the method's accuracy.

Haring and McMenamin constructed a regional model of Southern California using RAS, but it could not be tested, due to the lack of a survey table. "McMenamin has partially tested the method on Washington State data and found the results to be marginal at best " (McMenamin and Haring, 1974, p. 193). They also devised an alternate iterative procedure (H-M) but "[i]t is only applicable to the problem of estimating a regional table at one date based on a similar table from an earlier point in time." (1974, p. 194) This was tested, along with the RAS method and the naive method (no change from the original table) to estimate the 1967 Washington table from that of 1963. No adjustments were made for price changes. Results showed large errors in absolute percent deviation. By columnwise chi-square[8]

---

[8]
$$\chi^2 = \sum_i (a_{ij} - a^*_{ij})^2 / a_{ij} \ ;$$

this expression is also undefined when $a_{ij} = 0$ and $a^*_{ij} \neq 0$ and must be resolved arbitrarily.

calculations for coefficients, of the 27 sectors, only 10 in the naive model and 12 in the RAS and H-M were acceptable at the .05 level. Comparing their results to those of Schaffer and Chu, import averages and Types I and II multipliers were much better than Schaffer and Chu's location quotient, cross-industry quotient, supply-demand pool, and iterative procedure, but the authors are careful to point out that they began with a full-survey regional table, while Schaffer and Chu used a national table. McMenamin and Haring conclude their paper by stating that none of the methods they tested provide dramatic improvement over the naive model, that accuracy depends on the amount of structural change, and that it is worthwhile to estimate future regional tables if previous survey-based tables are available.

Malizia and Bond --

Malizia and Bond (1974) used the RAS method to update the 1963 Washington State table to 1967. They point out that the method should be restricted to the case where only one table exists, as Lecomber has shown, because other methods exploit the additional information of more tables. Basically, state Malizia and Bond, RAS deals with changes occurring when relative intersectoral relationships change between intermediate demand and final demand and between intermediate inputs and primary inputs. Its main assumption is that, given an initial change, the degree of change is proportional for each row and column. Their method of procedure was as follows. Making no price adjustments, interindustry flows were adjusted, using the 1963 table, so that each row sum gave 1967 intermediate purchases. Flows were then adjusted by 1967 intermediate sales and this process was repeated until residual errors were less than three percent (within five rounds).

The results of their test, comparing the 1967 table estimated from the 1963 survey table against the 1967 survey table, indicated high average coefficient error (104.7% - 132.7%) for various sector configurations, and relatively high information theory I-values (1.0 - 7.9). Further, they found large standard deviations (209.7% - 319.9%) and right skewness in the distributions. They state that "/_t_/hese data support the conclusion that the RAS method, used without additional exogenous information on interindustry flows for the projection year, is not powerful enough to generate satisfactory forecasts of interindustry coefficients." (p. 360) Comparing other aspects, Malizia and Bond found that forecasts of regional coefficients were more accurate than those of technical coefficients, aggregated tables were more accurate than disaggregated tables, and no relationship was found between degree of error by sector and per capita GNP, contrary to the claim of Chenery and Taylor, or between degree of error by sector and the location of the sector in a triangulated table.

Malizia and Bond also tested the predictive accuracy after aggregating to 18 sectors, using both RAS and actual tables. For total output and intermediate demand, "the predictions were quite correct;" for value-added, predictions were correct for the total and a few sectors. "However, in each case the synthetic table gave predictions that matched the accuracy of the actual table." (p. 362) Malizia and Bond conclude that the RAS method produces "large and theoretically unsystematic coefficient errors" so it cannot be recommended for forecasts of interindustry coefficients. "However, the method appears to be a useful way to adjust interindustry coefficients when short-term predictions of main aggregates are needed for regional planning and analysis." (p. 362)

Schaffer and Chu's Iterative Procedure

Schaffer and Chu (1964) developed another iterative procedure, the Regional Input-Output Table (R-I-O-T) Simulator. After using estimated regional output x_j for each industry to compute required inputs $r_{ij} = x_j a_{ij}$ and estimating regional final demand as a proportion of national final demand, regional sales for each industry are allocated initially using the national sales-distribution pattern. Then local sales are reallocated row-by-row using an iterative process until regional production and consumption are satisfied as much as possible. Exports are calculated as any positive differences between regional output and regionally required sales for an industry. The regional transactions matrix contains the positive differences between regional production requirements and available local output for each cell.

Comparisons of the Nonsurvey Methods

Schaffer and Chu --

Schaffer and Chu (1969) tested LQ, POLQ, CIQ, Pool technique, and R-I-O-T, concluding that "/ u /sing the 1958 transactions table for the United States and survey-determined industry outputs as program inputs, these simulations are comparable to the 1963 Washington input-output table and provide a limited test of acceptability." (p. 94) However, both Washington and national tables were aggregated to 23 industries before any calculations were performed. In addition, balancing for exports was required for LQ and POLQ. By the chi-square tests, LQ and CIQ were most successful, followed by R-I-O-T and the pool method.[9] The iterative procedure yielded the closest estimates of total imports into the state, while CIQ gave the best estimates of imports by industries alone and of exports. Imports to industries and exports are lowest for the iterative procedure, but since it attempts to allocate maximum possible local output to local trade, low estimates were expected. Mean income multipliers were high for all methods: for Type I multipliers, LQ gave results 21% higher, CIQ 38% higher; for Type II, LQ gave 47% higher, iterative 79% higher.

Schaffer and Chu conclude (p. 96) that the methods may be useful as supplements to survey studies, "/ b /ut it seems that, at the moment, there is still no acceptable substitute for a good survey-based study." Final notes based on their acknowledgedly limited test, include the observations that they avoided grossly mechanical computation and that their aggregation procedure ignored the Washington product mix. On an optimistic note they find that "it is surprising to us that our techniques yielded acceptable regional production coefficients for over 25 per cent of the industry categories," and if further tests yield evidence of consistent differences, they can be used as an adjusting factor.

---

[9] At $\alpha = .05$, the number of acceptable sectors was 6, 6, 5, and 4, respectively. (See Section V of this chapter.)

Morrison and Smith --

In a more comprehensive evaluation, Morrison and Smith (1974) estimated the 1968 Peterborough table from the national table using SLQ, POLQ, CIQ, RND, CMOD, RMOD, SDP, and RAS. The estimates were compared with a survey-based table by means of five tests: mean absolute difference, correlation coefficient, mean similarity index,[10] information content, and chi-square. Their results are presented qualitatively by the six techniques' ranking in each test. The RAS technique was best in all cases, with a coefficient of correlation of 0.501. This was not considered surprising, as the RAS method does use some survey material. The SLQ (employment rather than output was used) was second best, resulting as both the simplest and best of the totally nonsurvey approaches (but coefficient of correlation only equalled 0.160). Considering the other methods, POLQ was inferior to its unmodified version (SLQ) in all cases, but the modifications made to CIQ and RND (yielding CMOD and RMOD) did improve them. SDP gave the second best coefficient of correlation, but ranked poorly in the other tests. CIQ yielded the poorest results. A modification which resulted in estimates closer to the survey-based table was one in which zeroes were inserted in all principal diagonals of manufacturing and construction industries, as such interindustry transactions were theorized to be small. Morrison and Smith conclude, "[o]n the present evidence at least it would seem that nonsurvey methods can only produce an approximation of a full survey based table, and the adequacy of this approximation must depend on the objectives of the particular exercise." (p. 13)

---

[10]

$$S_{ij} = 1 - |a_{ij} - a^{*}_{ij}| \;/\; (a_{ij} + a^{*}_{ij}).$$

CHAPTER 4

A REVIEW OF ALTERNATIVE REGIONAL AND INTERREGIONAL MODELS

Following the rapid rise in energy costs, there has been both a public and private effort to locate new sources of energy. New energy development programs will obviously have an impact on the levels of economic activity and environmental quality, and hence should be evaluated in order to obtain the maximum benefit at the lowest cost from any program.

National models, either input-output or macro-econometric models, have been used to explain and project the effect of rising energy prices and energy supply constraints on the national economy.

Recently, a number of regional models have been developed to explain and project regional economic activity. Due to the complexity of the problem, and the not too infrequent lack of reliable data, most regional models are "linked" in some fashion to a national model. Examples of these models are the Shift-Share Models, Location-Agglomeration Models, Regional and Interregional Input-Output Models, and Regional Macro-Econometric Models. Each of these models was designed to explain and project regional economic activity given an exogenous change in public expenditure policy. At the original time of their formulation, however, these models were not designed for the purpose of analyzing the regional economic and environmental impact of alternative government energy policies.

In this chapter we will review and examine the different types of regional projection models. The existing regional forecasting models can be classified into four relatively distinct categories. These are:

I. The 'Shift-Share' Approach,
II. The 'Location-Agglomeration' Approach,
III. Regional and Interregional Input-Output Models, and
IV. The Regional Macro-Econometric Models.

Our discussion of the different types of models will be based on an examination of representative models. As an example of the first model we shall use the "OBERS Projections of Economic Activity in the U.S." The Curtis Harris, Multi-Regional, Multi-Industry Forecasting Model will serve as an example of the second type of model. The third, Polenske's Multi-Regional Input-Output Model, and fourth, Glickman's Regional Econometric Model of the Philadelphia SMSA, will be examined respectively.

Finally, we will provide a review of two other models describing the regional economies and energy use that are best described as "other models."

I. 'SHIFT-SHARE' APPROACH

OBERS Projections of Economic Activity in the U.S. (U.S. Water Resources Council, 1974, pp. 1-35)

Purpose

The OBERS projections were constructed to serve two purposes. These purposes are: (1) to estimate the economic effects of specified water and land resource constraints and of alternative programs for the development and management of the nation's water and land resources at the regional level; (2) to estimate regional demand for water and related land resources.

Geographic Region

Two geographic units were chosen for the OBERS study as the focal point for analysis. The BEA Economic Area was chosen in order to analyze the economic impact of a public investment project, such as water resource development. The Water Resources Council Water Resources Subarea was chosen to project agricultural production and land use.

Characteristics of the Model

Long-Run Forecasts --

The OBERS projections are based on long-run or secular trends and ignore the cyclical fluctuations which characterize the short-run path of the economy. The effects in terms of personal income, employment, and population changes are used to determine the effect on regional industrial and agricultural production.

Full Employment --

Nationally, reasonably full employment represented by a 4% unemployment rate is assumed to prevail at the points for which projections are made. Unemployment is disproportionately distributed regionally. Given this assumption, an increase in activity by a primary user of the proposed development project requires at least a partial offsetting reduction, or foregone expansion, elsewhere.

Regional Migration --

Workers will migrate to areas of economic opportunity and away from declining or slow growth areas. Consequently, regional earnings per worker, income per capita, and regional employment/population ratios will tend to move toward the national norm.

Technological Progress --

Continued technological progress and capital accumulation will support a growth in private output per man hour of 2.9% annually.

Regional Export Industries --
Most factors that have influenced historical shifts in regional 'export' industry location will continue into the future with varying degrees of intensity.

New Products --
New Products will be accommodated within the existing industrial classification system.

Variables Modeled into the OBERS Projections and Data Requirements

The effect on regional industrial and agricultural production in the OBERS system is calculated in terms of personal income, employment, and population changes.

Personal Income and Earnings in Constant Dollars --
Since there is no direct counterpart to GNP at the regional level, the personal income series and its earnings of persons component were used as an alternative to GNP. According to the OBERS projections, Volume 1, the choice rested on three considerations: (1) personal income has a close and comparatively constant relationship to GNP; (2) its regional location is clear; and (3) it could be measured from available data sources, and the methodology for preparing local area estimates of personal income had already been developed. Estimates of regional personal income were obtained at the local level by industry.

Industrial Output in Constant 1958 Dollars --
Projections of national industrial totals were similarly made for 37 SIC groupings. Projections of regional industrial output were obtained using a 'shift-share' estimation approach. The methodology for preparing the economic area projections will be described in the next section.

Employment and Population --
The employment series used in the projections of the national industrial structure, "persons engaged in production" is not available on a subnational basis. The only industrially complete employment series available for local areas is that from the decennial census of population. Therefore, the projected national employment "persons engaged in production" concept was converted to the census employment concept. Estimates of regional employment and population were obtained at the local level by industry.

Projection Methodology for Economic Areas

The OBERS projections were made in two major steps. First the national economy was projected in industrial detail. Secondly, these projected national totals were distributed regionally in accordance with projected trends in the regional distribution of economic activities. The projections, except for the determination of food and fiber, are calculated mainly from the supply side of the economy; therefore, there is an implicit assumption that sufficient demand will be generated by the private and public sectors in the production and distribution processes to maintain a full employment level of economic activity nationally.

The methodology for preparing the economic area projections utilizes four models. The first pertains to basic industries except agriculture and armed forces. The second projects the residentiary industries, with a modification of this model used to project property income, transfer payment and personal contributions to social insurance. The third relates to agriculture. The fourth model derives population from projected employment.

Basic Industries --

For the basic, or export-oriented, industries, the projection model was derived from the 'shift-share' technique for regional industrial analysis. Shift-share analysis is designed to discern regional departures from national industrial growth rates. In its simplest form, this technique distinguishes a proportional growth element and a differential growth element between a region and the nation.

$$E_{ij}^{t} = (E_{io}^{t} / E_{io}^{x})\ E_{ij}^{x} + C_{ij}^{x-t}$$

where the subscripts i, j refers to the ith industry and jth region, and the subscript o refers to a summation. Superscripts t, x refer to the projected time point and the base point respectively. Finally C_{ij}^{x-t} refers to the difference between the hypothetical level attained at the national growth rate of the industry over the period x to t and the regional level actually attained in the industry over the same period.

The C_{ij} term was projected by fitting a curve to each region's percent of the national total income and employment (separately) for the selected years for which data was available. The curve was then extended into the future and the values of the region's projected percentage in the target years read from the curve. This technique yields a trend extension of a region's historic percent of the national total of earnings and employment in a given industry. It was accomplished by using linear regression analysis and applying substantial judgment to dispose of erratic or cyclical observations in the historical time series including apparent discrepancies between the earnings and employment series.

Residentiary Industries --

Local-service or residentiary industries primarily serve local businesses and households with commodities and services which do not enter into interregional trade in substantial amounts. The outputs of these industries, therefore, are determined by the size of population and personal income in each region.

Agriculture --

Projections of agricultural production and land use did not go through the economic area stage, but rather went through the projection methodology for water resources regions, subareas and states. The projections are based on the same framework of assumptions as the other industrial sectors, and on the extension of historical trends and the use of a land availability check to ensure the adequacy of resources within the regions to produce the projected output.

Projections of agricultural output for the water resources regions and subareas were developed through a three-stage process: First, the projected national output was distributed among states. Second, the state totals were disaggregated to subareas and regions. Finally, the output projections in percentage terms were converted to quantity and value estimates.

Population --

It is assumed regional variations in the rate of natural increase in population in an area are small, therefore, local area population projections were based on interarea migration. Assuming the major motivating factor in migration is economic opportunity, area population for the most part was projected as a function of area employment. To account for non-economic causes of interarea migration, the population was divided into three categories: the pre-labor pool (ages 0-14), the labor pool (ages 15-64), and the post-labor pool (ages 65 and over). Population in the labor pool was projected as a function of area employment. The pre-labor pool population was projected as a function of the total labor pool, while the post-labor pool population was projected in terms of population aged 55 and over at the preceding decade using the following formula:

$$\frac{P^{65+}_{t}}{P^{55+}_{t}} = \frac{P^{65+}_{t-1}}{P^{55+}_{t-1}}$$

This formulation holds constant the relationship between the 55 and over age group at a given time and the 65 and over age group ten years later.

II. 'LOCATION-AGGLOMERATION' APPROACH[1]

C. Harris' Multi-Regional, Multi-Industry Forecasting Model (Harris, 1973)

Purpose

The second type of model we considered in our investigation was the Curits Harris Multi-Regional, Multi-Industry Forecasting Model. This model is designed to predict the short-run impact and project the long-run effect of an exogenous change in public expenditures. In order to maintain consistency between regional and national forecasts, the Harris model, like the OBERS system, projects regional shares which are then applied to the projected national values to obtain regional values.

---

[1] We have classified the Harris Model as the "location-agglomeration" approach due to the multitude of variables used in his equations to explain economic activity.

Geographic Region

The geographic unit chosen as the focal point for analysis is the county, though in a later study (Harris 1974) Harris aggregates the 3111 counties into the 173 OBE Economic Areas.

Characteristics of the Model

The following assumptions characterize the Harris model.

(1) The estimates of national coefficients can be applied uniformly at the regional level.

(2) The Harris equations assume uniform production and price functions among regions (Miernyk 1973, p. 281).

(3) Changing shares in regional production resulting from changes in transportation costs also reflect the dynamic character of the model. These changing shares underlie the shifting location of regional industries.

(4) There are no abrupt changes in any regional economy, but rather the changes that do take place result from the gradual adoption of innovative technology.

Variables in the Harris Model and Data Requirements

The Harris equations are equipped to forecast industrial and construction output, employment, population, personal income, personal consumption, government and defense expenditures, gross investment for a number of industries, and marginal transport costs.

All variables, except the marginal transport costs, were estimated using OLS regression analysis. Values for the transport costs were obtained from a linear programming solution to a transportation cost minimization problem.

Projection Methodology

The model is recursive in that supply and demand data in year t are used to forecast variables in the year t+1. The forecasts are then used as data to make forecasts for the year t+2, and so on. In specifying the model, Harris used a modification of a procedure referred to by Theil as "extending the set of explanatory variables." According to Theil, this procedure involves establishing a critical set of explanatory variables in the initial regression; then additional independent variables are introduced in order of their theoretical importance in a step-wise regression program. Variables are added until a coefficient becomes insignificant, at which point the procedure stops. Harris uses this approach: the least squares regression routine was set up to preserve the theoretical order of variables entering the equation, but was not terminated when one variable became insignificant.

In order for a variable to be retained in a particular equation of the model it had to pass two tests in addition to the standard "t" test of statistical significance. First, the sign on the coefficient had to have the postulated theoretical sign, the entering variable itself could not have severe multi-collinearity with other variables in the equation. The decision rule for dropping a multi-collinear variable from an equation is that a variable was rejected if the coefficient of determination (R^2) of the independent variable to be entered with those independent variables already in the equation exceeded 0.70. The decision rule used for the test of significance was that the 't' ratio had to be greater than one (the standard error could not exceed the coefficient).

The forecasting model starts by forecasting output by industry in each area for the first year after the base year (1965). The output or change in output of each industry sector is explained by the marginal costs or prices that firms face in each location. After output has been determined, then payrolls, employment, population and personal income are derived. The final demand sectors are also forecast. Consumption and governmental expenditures are related to income; investment is related to output; and, foreign exports are determined exogenously.

The parameters for each equation used in the forecasts were estimated using Ordinary Least Squares in a cross-sectional regression of approximately 3,111 observations. 1965 was selected as a base year because it was the only year for which a complete set of data for all counties was available. In some equations some of the observations were eliminated if the value of the dependent variable was zero. In the location equations, observations were eliminated in some industries because of other data problems. Data for the marginal cost variables were obtained using a linear programming transportation algorithm.

The Harris model also uses dummy variables for each of the regional observations to capture some of the variation in the dependent variable not fully explained by the explanatory variables. Harris assumes there is some unique constant in every county that is not being explained by a particular equation. Therefore, there is a separate, though parallel regression line for each county.

III. REGIONAL AND INTER-REGIONAL INPUT-OUTPUT MODELS

K. Polenske's Multi-Regional Input-Output Model (Polenske, 1973, pp. 1-37)

Purpose

A third model considered is K. Polenske's Multi-Regional Input-Output Model (MRIO). This model, like Harris' system of equations, is designed to analyze the short-run impact and predict the long-run effect of an exogenous change in government expenditures on a region's economy. MRIO goes one step further, however, in that it is also designed to measure the interregional repercussions resulting from a change in government policy. The interregional economic repercussions are examined in terms of the changed interregional trade flows.

Geographic Region

The focal point of regional economic analysis in MRIO is the state. The complete regional and interregional accounting system consists of an I-O table for each of the states and the District of Columbia and 2550 trade flow matrices illustrating the commodity flows between the regions.

Characteristics of the Model

Unlike the previous two types of models, the Polenske model does not assume uniform regional production functions (except in the trade and service sectors).

Technical Coefficients of Production --

Separate technical coefficients were calculated, by state, for agriculture, mining and new construction. In the manufacturing sectors, national coefficients were adjusted to allow for differences in product mix. Unadjusted national coefficients were used for the trade and service sectors.

Final Demand and Trade Patterns --

At any given point in time final demand and the coefficients are assumed exogenous. With variation in final demand and the coefficients over time, the Polenske model allows for regional shifts in production, consumption, and trade.

Price Sensitivity of Output --

Given the nature of the Polenske model, output is not responsive to changes in relative prices. Regional production may vary only as the result of a change in demand for that output.

Variables in the MRIO Model and Data Requirements

The complete multi-regional Input-Output accounting system consists of a regional I-O matrix for each of the 51 regions and 2550 trade flow matrices reflecting the interregional commodity flows between regions. Therefore for the complete multi-regional I-O accounting system 2601 regional blocks of data are required. The regional I-O matrices are designed to give purchases by the location of the purchaser, while the trade flow matrices are designed to give the total inputs of a commodity i required by industry j located in region g from all regions (including region g);

e.g., $X_{ij}^{g} = \sum_{h=1}^{51} X_{ij}^{hg} \quad \forall \; i,j,g$

where h specifies the region in which the good was produced.

The regional I-O table consists of a matrix of inter-industry flows $||a_{ij}^{g}||$, a vector of final demand, and a vector of value added. The technical coefficient, a_{ij}^{g}, is defined in value terms as the "total amount of the input of industry i in region g required per unit of output of industry j. Since each element of the regional table gives the purchase of a commodity without regard to the location of production, each row shows the total distribution of a commodity to the intermediate and final users within that region. The rows specify the producing industries but do not designate the region in which the goods were produced. Similarly, each column of the table shows the total purchases of goods, services and value added by the particular intermediate or final user located in that region.

Usually, in a closed economy, the sums of corresponding rows and columns of an input-output table will be equal. This, however, is not true when dealing with relatively open regional economies. The discrepency between the sums of corresponding rows and sums is due to the volume of interstate trade flows. If a specific commodity is not traded, then it can be expected that the sum of the corresponding row and column will be equal. Using interregional I-O tables, each total purchase made by an intermediate or final user in a particular region can be traced back to its source. In the interregional trade tables, each row sum indicates the total production of each commodity in each state, while each column gives the total purchases of each commodity by each region.

For the MRIO analysis, detailed state input requirements for 1963 were assembled for three large sectors of the economy: agriculture, mining, and construction. Most of the research effort on regional differences in technology was concentrated on these sectors because locational factors are likely to cause significant state-to-state variations in their input requirements. For the remaining industries, namely the manufacturing and service industries, either variations in the state coefficients were obtained through the product mix method,[2] or for few of those industries, where the components of the industry could not be separated, the national coefficients were used directly.

Methodology

In estimating the regional technical coefficients, K. Polenske made use of the national primary and secondary matrices. The primary product matrix is a product-by-establishment matrix, in which the rows show the total amount of a given product distributed by each establishment regardless of which establishment produced the product and the elements in the columns show the total amount of a particular good purchased by establishments to produce the total output of the establishments -- their primary and secondary products.

The secondary product matrix is an establishment-by-product matrix that provides information on the composition of secondary-product output for each industry. The elements in each row give the value of each of the secondary commodities produced by that establishment according to the industry of which they are the primary commodity. The elements in each column, on the other hand, give the value of commodities of the given establishment that are produced by other industries. By definition, all diagonal elements are zero, since each element s_{ii} represents the quantity of goods of industry i that is produced by industry i, which cannot be secondary products of industry i.

---

[2] Product Mix Method: To obtain state inputs, the national direct primary input coefficients for the manufacturing and service industries are weighted by the state establishment outputs. The resulting regional input requirements for a given industry varied from state to state, reflecting regional variations in the composition of goods produced within different states.

Since regional secondary product data were not readily available, all of the regional technology data assembled for the MRIO study were reconciled with the OBE 1963 national primary-product matrix, while the regional secondary products were estimated using the national primary and secondary matrices and state outputs as a basis.

Having estimated the elements in the regional secondary product matrices they were combined into a single row and column and inserted as an extra row and column along with those of the regional primary matrices to obtain the total transactions matrix $||a^g_{ij}||$ for each region.

IV. REGIONAL MACRO-ECONOMETRIC MODELS

N. Glickman's "Philadelphia SMSA Regional Econometric Model" (Glickman, 1971)

Purpose

The fourth general type of model we considered is a regional macro-econometric model. This type of model, theoretically, assumes one output per region, but Glickman relaxes this assumption in order to forecast output by various manufacturing and non-manufacturing industries.

The Glickman model has two principal purposes. First, it can be used for macro-econometric forecasting analysis of regional economies; and second, it can be used for multiplier analysis of policy-related phenomena.

Geographic Region

The Glickman model was constructed to analyze the regional economy of the Philadelphia SMSA.

Characteristics of the Model

The Glickman Model was designed as part of a satellite system to be attached to a system for the national economy as a whole. Unlike the Polenske model, it is a "top-down" rather than "bottom-up" model, implying the national economy 'drives' the regional economies. This approach is taken because of the lack of regional data.

Due to the data constraints only annual data can be used in the regressions, and even then relatively few economic variables are used.

Variables Included in the Model and Data Requirements

The Philadelphia SMSA Regional Econometric Model consists of fourteen blocks of equations, aimed at providing a fairly comprehensive view of the region's economy. These blocks are categorized as follows:

(1) manufacturing output
(2) non-manufacturing output
(3) manufacturing employment
(4) non-manufacturing employment
(5) wages, prices, and income
(6) Federal and local government
(7) manufacturing investment
(8) demographic
(9) retail sales
(10) banking
(11) the City of Philadelphia Accounts
(12) the suburban subregion accounts
(13) consumption
(14) a quarterly submodel

Of these equations, forty-seven equations summarize the important private and public sector activity in the City. Given the variables determined there and for the region as a whole, a residual block for the 'suburban' portion of the region is calculated. Consequently, three separate area units are considered; the City of Philadelphia, the suburbs, and the entire SMSA.

In order to specify many of the model's equations, 'a priori' information was used from the Philadelphia input-output table. For example, the import and export columns of the table were used to delineate local from export-oriented industries. The specification of equations in the manufacturing output and employment blocks was also facilitated by noting the important interindustry relationships in the transactions matrix.

Methods of Estimation

The blocks in the Philadelphia model are related primarily in a simultaneous rather than a recursive manner. To the extent the equations are recursive, they were estimated by Ordinary Least Squares; otherwise, the equations were estimated by Two-Stage Least Squares -- Principal Components and Iterative Instrumental Variables.

> "In the Philadelphia model, national market-oriented output is related to national variables such as GNP and national trends in those industries and to local interindustry partners. The model is designed to be 'plugged in' to the Wharton Annual and Industry Long-Term Model of the United States, taking as exogenous many of the variables forecasted by the Wharton model such as Gross National Product and U.S. output and wages in various industries. On the other hand, output which serves a local market is a function of Philadelphia's personal income and other measures of the local economy. Thus, a modified economic base method

is employed. Employment by industrial sector is related to sector output as in an inverse production function.

There is a rather large Federal and local government section. In addition to determining Federal personal income tax accruals from the Philadelphia region, the model estimates tax revenues for both municipalities and school districts; certain governmental expenditure variables are endogenously determined. A large number of policy instruments on the national and local level are contained in the model, including inter-governmental revenues, the Federal personal income tax rate, defense spending, Philadelphia City tax rates and Federal government employment in the region. These policy variables can be manipulated to examine the impact of government policy changes in the region: some of these simulation experiments are described in pages below.

In the income block, wages by sector are related to national wages and the local unemployment rate. Proprietors' income is related to local retail sales and population and national trends in proprietors' income. Transfer payments are determined by the corporate bond rate (representing private transfers) and total U.S. transfer payments (proxying public transfers). Personal income is the sum of wages and all non-wage income. Consumer prices are a function of local unit labor costs, national consumer prices and regional profits.

Retail sales for the region are related to personal income and the employment block helps determine the demographic block; the latter includes the unemployment rate, among other variables. The output and personal income block equations impinge on the banking sector where demand and time deposits for the local Federal Reserve district are determined.

A small quarterly submodel predicts important regional variables such as manufacturing and non-manufacturing employment, unemployment and the consumer price index. Thus, there is a mixture of annual and quarterly data within the model.

In addition to time-series data, cross-section information is employed in the construction of the model. Eleven components of household consumption expenditures are estimated from a 1960-1961 BLS sample survey." (Klein and Glickman, 1977)

V. SUMMARY OF FIRST FOUR TYPES OF MODELS

We have reviewed a variety of existing regional models with the objective of determining which of these models can be incorporated, with or without modification as the Regional Module of SEAS. Due to the complexity of the problem, and the not too infrequent lack of reliable data, most regional models are 'linked' in some fashion to a national model. Examples of these models are the (1) Shift-Share, and (2) Location-Agglomeration Models, (3) Regional and Interregional Input-Output Models, and (4) Regional Macro-Econometric Models. Each of these models is designed to explain and project regional economic activity. At the original time of their formulation, however, these models were not designed for the purpose of analyzing the economic and environmental impact of alternative government energy policies.

Existing Regional Forecasting Models

The existing regional forecasting models can be classified into four relatively distinct categories and these are: 1) the OBERS "shift-share" type model; 2) the Harris "Location-Agglomeration" type model; 3) the Klein-Glickman type Macro-regional econometric models; and 4) the Polenske MRIO model. These categories are not, however, entirely distinct since some of the models utilize a number of different techniques to conduct forecasting analysis. For example, the Harris model and the OBERS projections combine regression analysis with the regional shares, and I-O approaches to determine the effect of alternative transportation systems and water resource development on a regional economy, respectively. MRIO is strictly an input-output technique developed to project regional and interregional economic activity, and the Glickman model is strictly an econometric type model using regression analysis to project economic activity in the Philadelphia region.

Energy and the Environment

None of the models, however, contains explicit formulations to project the environmental impact and the effect on energy demand of alternative energy scenarios. The Glickman model and MRIO are 'pure' regional economic impact models. The OBERS model tries to project the environmental impact of water resource development on land use, and the Harris model introduces environmental quality variables to explain regional variation in output.

The Multiregional Economic & Energy Demand Model (MREED) seeks to integrate energy and environmental impact variables with some of the better features of the models studied. (The MREED model is presented in Chapter 5.)

Structure of the Models

All of the models, except MRIO, consider the national economy as the 'driving force behind the regional economies.' The OBERS projections and the Harris model project regional shares of national values -- output, employment, income -- into the future. Changing shares over time reflect the dynamic character of the model and the shifting location of regional industries. The Glickman regional model of the Philadelphia SMSA was designed to be part of a satellite system with the national economy in the center of the system. MRIO, on the other hand, was designed to estimate each region's level of output, employment, income and expenditures, and then obtain the national totals as the sum of all the regional values.

The MREED model will serve as the Regional Module of SEAS; therefore, it will project, along the OBERS and Harris lines, regional shares of national totals.

Geographical Unit of Analysis

To project economic activity levels OBERS bases its analysis on the BEA Economic Area. The Harris model has also been applied to the 173 Economic Areas, but was initially intended for the county subdivision. The MRIO region is the state, a political unit. The Glickman model is based on the Philadelphia SMSA. Of these modular regions, the MRIO state is probably the most open since SMSA's extend beyond political boundaries. The BEA economic areas are the most self-sufficient in the employment sense since they usually contain labor's place of residence and place of employment in the same area.

The MREED model, partially for this reason, will be based on the BEA Economic Areas.

VI. OTHER MODELS

THE FEDERAL ENERGY ADMINISTRATION'S READ MODEL[3]

As must be expected with models currently under development, a comprehensive evaluative description of the Federal Energy Administration's Regional Energy, Activity and Demographic Model (READ) was not available for external review. As such, the internal workings of READ could not be subjected to as great a detailed examination as would have been preferred. However, a partial description of its structure (consisting of general flow charts and two memoranda) did permit a review of some of the model's conceptual underpinnings and linkages.

General Structure of the READ Model

The Federal Energy Administration's Regional Energy, Activity and Demographic Model (READ) generates forecasts of county-level output, by industrial sector, based on exogenous energy specifications. The county forecasts are then aggregated to a regional level for a structural demand model which assigns the forecasted regional outputs to the least-cost technological process, determined by the structural model. In this way, the model's projected demand for energy use is partially a function of the READ forecasts.

---

[3] The document on which this review is based came to us from ERDA at one of the SEAS design meetings.

From the model structure depicted in Figure 4-1, data bases from various federal government agencies are used as inputs for such types of parameters as commerce data, demographic information, employment, income, government, and weather. Direct inputs to READ include regional energy prices generated by Project Independence (PIES), national macroeconomic, employment, income, and population forecasts, the previous year's regional supply, demand, population, and employment, and governmental structure.

READ then is used to project energy demand by sector by region. Location equations estimate regional output which leads to agricultural, industrial, and commercial energy demand. Demographic estimates are used to determine energy demanded by residential and commutation activity, with the latter in turn influencing transportation energy demand, subject also to transportation requirements and shadow prices. Energy required by government is a function of government expenditures.

Industrial Structural Demand

The READ model also will be used to forecast regional industrial capacity, and thus estimates of annual equipment expenditure and construction purchases by industry. These forecasts interface with the structural energy demand equations in the following manner (the model will be utilized for the residential (Hirst Model), commercial, and industrial sectors, but we will discuss only the industrial sectors).

The industrial structural demand for an industry is calaculated as:

$$Q_{ijr} = f(\text{Capacity}_{r}, \text{Technology Penetration}_{j,r}, \text{Short-Run Utilization}_{j,r}, \text{Fuel Usage}_{i,j,r})$$

for fuel i, technology j, region r. Initially, structural models are being developed for the ten highest energy-using industries. The total capacity of the industry in region r is the maximum economically feasible output level with existing capital stock. (The measurement of this parameter is considered highly feasible, because the energy-intensive industries are primarily continuous process ones.) The capacity is not specific to technology (e.g., "steel mill output").

The technological penetration parameter uses the alternative energy-using industrial process and their individual capacities. A least-cost basis is used to assign technological process. The Demand Analysis Division (DAD) is in the process of determining the most probable future technological processes in use in 1990, including estimates of technological shifts over time. The READ equipment expenditure forecasts and the technological projections will enable DAD to forecast levels of technological penetration. (The residential and commercial sectors will be driven

Reproduced from
best available copy.

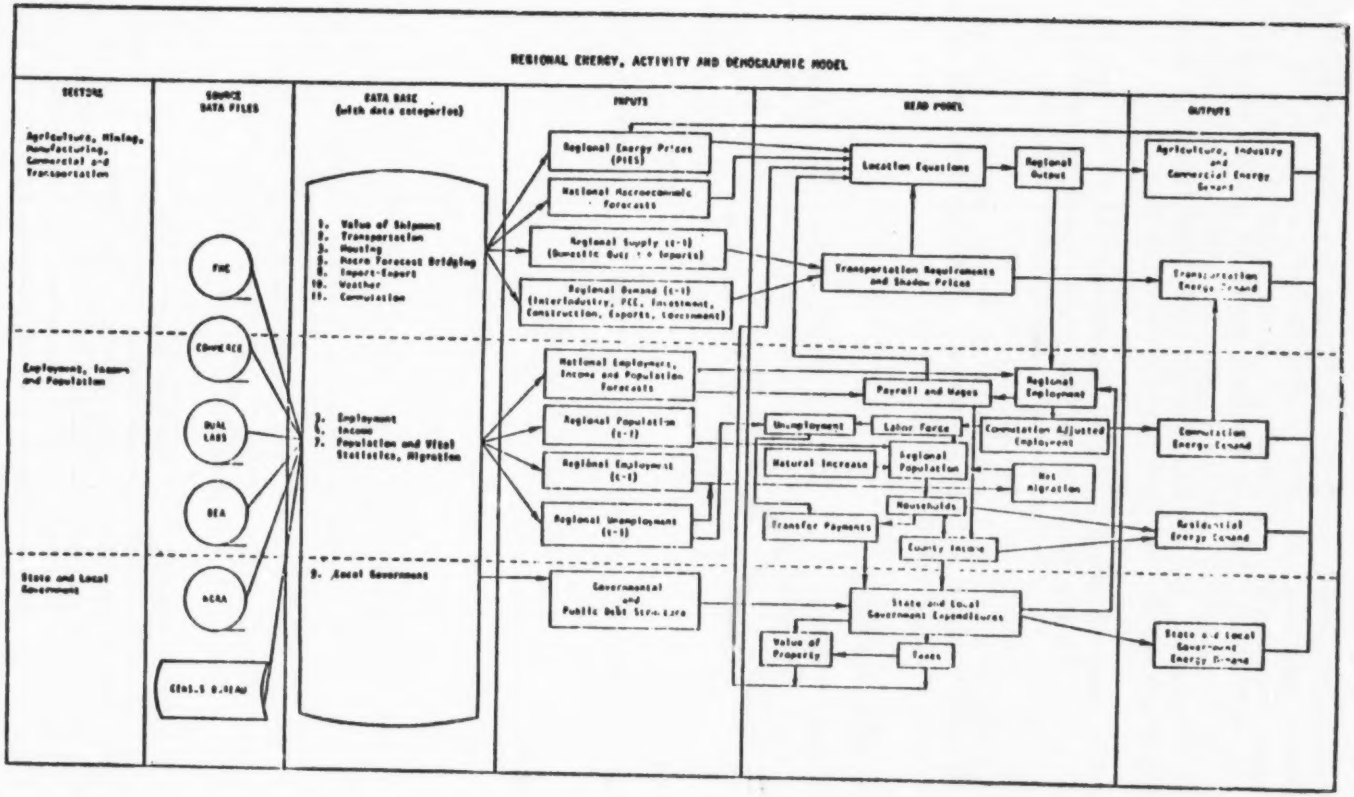

REGIONAL ENERGY, ACTIVITY AND DEMOGRAPHIC MODEL
SECTORS
SOURCE DATA FILES
DATA BASE (with data categories)
INPUTS
READ MODEL
OUTPUTS
Agriculture, Mining, Manufacturing, Commercial and Transportation
Employment, Income and Population
State and Local Government
FPC
COMMERCE
DUAL LABS
BEA
NOAA
CENSUS BUREAU
1. Value of Shipment
2. Transportation
3. Housing
5. Macro Forecast Bridging
8. Import-Export
10. Weather
11. Commutation
4. Employment
6. Income
7. Population and Vital Statistics, Migration
9. Local Government
Regional Energy Prices (PIES)
National Macroeconomic Forecasts
Regional Supply (t-1) (Domestic Output + Imports)
Regional Demand (t-1) (Interindustry, PCE, Investment, Construction, Exports, Government)
National Employment, Income and Population Forecasts
Regional Population (t-1)
Regional Employment (t-1)
Regional Unemployment (t-1)
Governmental and Public Debt Structure
Location Equations
Regional Output
Transportation Requirements and Shadow Prices
Payroll and Wages
Regional Employment
Unemployment
Labor Force
Commutation Adjusted Employment
Natural Increase
Regional Population
Net Migration
Households
Transfer Payments
County Income
State and Local Government Expenditures
Value of Property
Taxes
Agriculture, Industry and Commercial Energy Demand
Transportation Energy Demand
Commutation Energy Demand
Residential Energy Demand
State and Local Government Energy Demand

Figure 4-1. Flowchart of READ model. (source: personal communication)

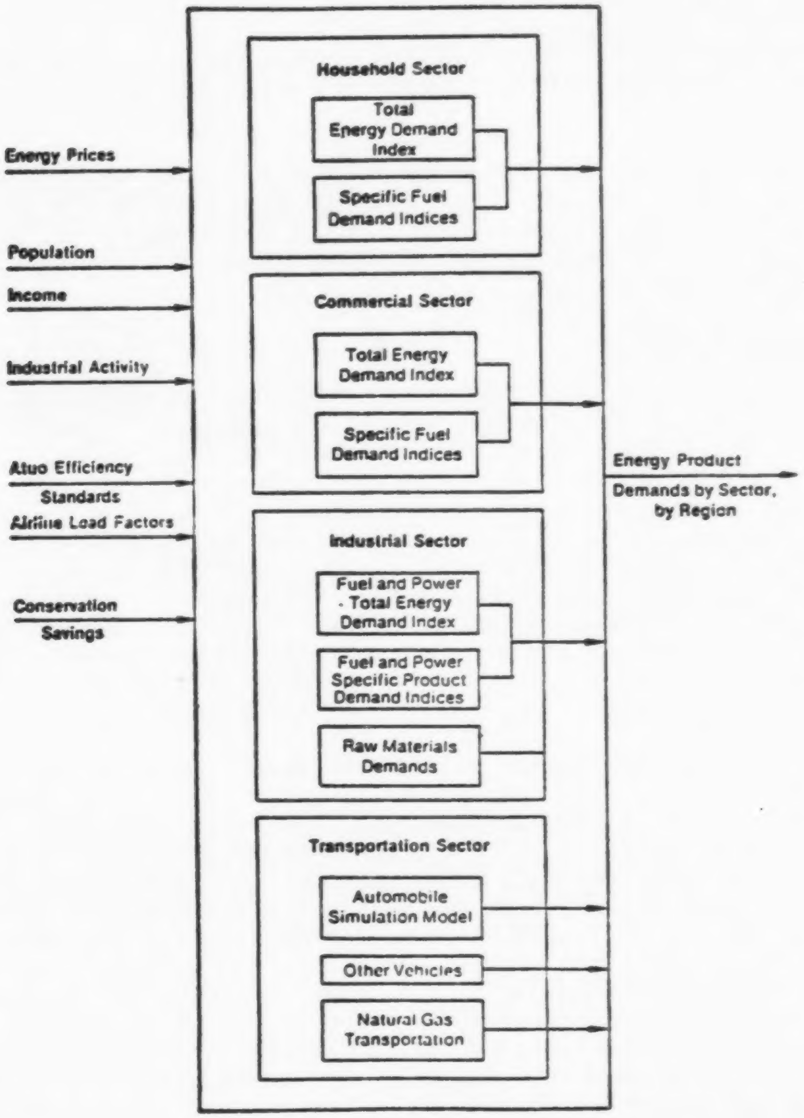

Energy Prices
Population
Income
Industrial Activity
Atuo Efficiency Standards
Airline Load Factors
Conservation Savings
Household Sector
Total Energy Demand Index
Specific Fuel Demand Indices
Commercial Sector
Total Energy Demand Index
Specific Fuel Demand Indices
Industrial Sector
Fuel and Power - Total Energy Demand Index
Fuel and Power Specific Product Demand Indices
Raw Materials Demands
Transportation Sector
Automobile Simulation Model
Other Vehicles
Natural Gas Transportation
Energy Product Demands by Sector, by Region

Figure 4-2. Energy demand in READ.
(source: personal communication)

as well by READ's forecasts of population and residential and commercial construction expenditures.) An example of the model's hypothetical technology penetration for the r^{th} region with technology j is "electric blast furnace."

The short-run utilization variable indicates the fraction of process capacity currently being used in the production process. Industrial output by region is obtained by multiplying utilization by capacity. Short-run utilization would typically be a function of fuel prices, output prices, and possible prices of other factors. For the j^{th} technology in region r, the parameter can include the number of shifts and number of hours.

The component of fuel demand in the equation allows technological transportation from a particular process's level of output to its required amount of fuel i.

While an alternative source of capacity and output forecasts can be used instead of READ, two advantages of using READ to drive the structural demand models are that (1) both models share a common data base consistent with PIES, and (2) the industrial location and population sectors of READ are energy-dependent.

Other Components of READ

As previously explained, little description of the READ model's structure was available for review, other than Figures 4-1 and 4-2 (attached) which are self-explanatory. Various exogenous parameters serve as inputs to the model, including demographic data, outside agency data and forecasts, and intra-FEA data and forecasts (or scenarios). The absence of almost any detailed description of the source and acceptability of data, and the structure of the READ model, prevents analysis of its potential usefulness (although the presence, in Figure 4-1 of a feedback mechanism for output energy demands to be incorporated in new energy price calculations, does provide evidence of one potentially favorable structural aspect). The exception of the construction sector, for which a six-page memorandum was available; does allow some detailed analysis.

In the construction sector, the straightforward computations of capital stocks do not leave much room for argument, except for the fact that they incorporate some rather gross assumptions. The Dodge Construction Potentials data must be adjusted for "many" differences to conform to National Income Accounts, a task found "impossible," other than by simply scaling county data to aggregate National Income controls. Aggregate National Income Accounts deflators will be used to convert county construction data to constant dollars. Farm construction and oil and gas well drilling and exploration data are not available from Dodge, and potential sources are currently being investigated. State-level input-output multipliers for all industries (including construction) are available in sufficient detail for some construction sectors. However, others must use national coefficients. Capital stocks are unavailable at the regional level, and "(even) national estimates of capital stocks are far from ideal," although estimates "by considerable industrial detail...(and)...by types of plant and equipment" "can be constructed." National capital-output ratios or capital-labor ratios will be applied to county output or employment. National discard rates as well (historical and projected) will be applied. The lack of much required regional data, at least in the construction sector, does cast doubt on the confidence which can be placed on regional estimates.

Summary of READ Model

From the information on READ available to us, the model appears to have a reasonable structure. However, several potential problems must be confronted.

First, one structural aspect which can be questioned upon examination of the flowcharts is the projection of output on a county level. The potential error inherent in attempting such micro-level projections based on national forecasts raises questions about the accuracy of such estimates. Further, such errors would be compounded in any aggregation up to a larger regional level.

Second, as evidenced by the problems in obtaining regional data for the construction sector, it would appear that estimation and calibration at a county level would be subject to great data limitations in all areas. This may result, as in the construction sector case, in unfortunate reliance on national parameters in many cases.

Third, if the projection of regional output depends on only capacity and utilization (based on prices), as it appears, the regional allocation process itself is suspect.

While other questions could be raised about the structure of READ if more detailed information would be available (particularly the flowchart's "location equations" which determine regional output), the ability of READ to project confidently regional economic, environmental and energy outputs would appear suspect, at least in its present working-stage form.

OAK RIDGE NATIONAL LABORATORY'S MULTIREGION (OLSEN, 1976)

While other models examined in this study generate projections of regional economic output, the stated purpose of MULTIREGION is to handle "core" requirements of regional decision makers and energy planners -- "basically accurate and sufficiently detailed forecasts of regional population and employment." (Olsen, p. 1) Developed by the Regional and Urban Studies Section of the Energy Division of Oak Ridge National Laboratory, MULTIREGION is designed only for these "core" requirements, while traditional methods are to be used to obtain other types of projections. As such, the model is inappropriate, or at least incomplete, for incorporation into SEAS as a regionalization component. However, we will describe its structure as an alternative approach to regional modeling.

Overview of MULTIREGION

MULTIREGION attempts to combine four forecasting methods - trend projection, cohort-survival analysis, export-base analysis, and shift-share analysis - so that divergent forecasts typical in their separate applications can be avoided.

By viewing Bureau of Economic Analysis (BEA) economic areas as labor markets, regional and interregional supply and demand processes can be modeled. Progressing in five-year time steps, for each BEA economic area, computations are made of:

(1) trial population (last period population + births - deaths + net migration)
(2) trial labor supplies (population times labor participation rates)
(3) trial labor demand (sum of forecasted employment)
(4) trial labor market conditions (from labor supply and demand)
(5) final labor market conditions (reiterating (1) - (4)).

Factors in the Model

Little consideration is given to economic influences in the demographic effects projected by MULTIREGION. Mortality and fertility vary over regions, but without reference to labor market conditions, while migration is affected by labor market tightness, interregional accessibility, and population density. Finally, labor force participation rates depend on labor market tightness and female participation in the labor force (a reflection of the regional industry mix).

The employment sector of the model includes three components: manufacturing, service, and natural-resource-based employment.

The natural-resource-based sectors use a shift-and-share framework to exogenously treat activity levels of those sectors "fairly exogenously." In this way, labor market conditions do not affect these sectors' employment.

Employment in the manufacturing sectors takes into account initial conditions (past locations of employment) and interregional market size as the two most important influences on industrial locational change.

Regional service employment projections use initial conditions, total employment growth, and to a lesser extent interregional market size, population density, and labor market tightness.

Outputs of the Model

The outputs of MULTIREGION are intended to resemble those of the OBERS in both form and potential applications. For both the nation and each of the 173 individual BEA economic areas, historic and projected values are given for employment by industry sectors (37), age and sex cohorts (32), and labor force cohorts by age and sex (16). Rates of growth and summary measures such as interregional market potentials, industry location quotients, coefficients of specialization, and shift-and-share components are also generated as output by MULTIREGION.

Summary and Conclusions

By design serving only a limited purpose, the demographic and employment projections estimated by MULTIREGION are inappropriate for the regionalization module of SEAS. The existing weakness of REGION can be attributed to the use of employment to allocate output. Further, alternate development or policy scenarios cannot be tested in a model which deals essentially as basically one of demographics and only in implicit economic interactions. Finally, the dynamic effects of potential changes in the economy (e.g. Western coal development, air emissions regulations, or drastic energy price changes) cannot be reflected in a model relying heavily on past trends and the "inertia" of "initial conditions." For these reasons, MULTIREGION would be an inappropriate choice for incorporation into SEAS.

CHAPTER 5

A MULTI-REGIONAL MODEL OF REGIONAL ECONOMY AND ENERGY DEMAND (MREED)

I. INTRODUCTION

In an affluent, spatially integrated economy such as the U.S., the regional economies are intimately linked to one another and to the national economy. The size and nature of the regional economy are largely determined in the national markets and by the spatial configuration of activities in national space.

What constitutes the national economy at any point in time is rooted in the determinants of final demand -- income distribution and consumer preferences, and the contemporary organization and technology of production. As the latter change, so do the size and composition of the national economy, and its spatial manifestation -- the regional economies. Location theorists have identified several distinct phases, each of which is characterized by the importance of different locational factors, in the history of the spatial evolution of the American economy (Chinitz 1960, C.D. Harris 1954, Ullman 1958, Perloff and Wingo 1961, Berry and Neils 1969). These locational factors stem from the multiple influences of natural resources and technology (of transportation) in the earlier periods and amenity and environment in the contemporary period.

Following the tenets of classical location theory, Chinitz recognizes two historical phases in the distribution of production in the U.S. largely by the structural changes in transportation. In the 19th Century, the rail network which reduced the cost of long distance travel relative to short haul led to increased concentration of manufacturing industries in the U.S. Early in the 20th Century, technological developments that reduced the cost of short haul relative to long haul (e.g., truck transport) led to a second phase of decentralization of industry.

The 20th Century has seen an increasing role of the market as a determinant of location of industries in the nation. Harris and Ullman have documented this by showing the juxtaposition of population and economic activity in the industrial 'heartland' or the manufacturing belt extending westwards from New York, in the area bounded by the Pennsylvania coalfields, Lake Superior ores and the capital, entrepreneurial experience and the engineering trades of the North East (roughly Middle Atlantic, New England, and Great Lakes Regions). Berry and Neils have shown that this heartland early became the center of heavy industry in the country but has since remained the center of national demand, determining patterns of market accessibility. Perloff and Wingo show further the historical tendency for fabricating industries to concentrate in the heartland while first-stage resource users or processing industries tend to dominate in the hinterland. Perloff and Wingo suggest a spatial pattern of the national economy as

> a great heartland nucleation of industry and the national market, the focus of large-scale nation-serving industry...(Outside the heartland) are the resource-dominant regional hinterlands specializing in the production of resource and intermediate inputs for which the heartland has need...[1]

Perloff and Berry and Neils have noted an emerging factor in regional economic growth in the second half of this century. This is the amenity factor and environmental quality the so-called "new natural resources that count." In the contemporary affluent urbanized age, those new resources are clean air and water, open space, and other amenities. Berry and Neils trace the emerging relationships between these environmental amenity features and regional growth outside the heartland in the South and Southwest - the so-called Sun Belt. Further, rising energy prices in the last decade have begun to affect regional purchasing power and energy induced structure of transportation rates (Miernyk 1976(a), Miernyk 1976(b), Giarratoni and Socher 1977). Since the energy resources are largely in the national hinterland, significant regional economic and locational effects are likely to ensue.

The regional models developed so far have not addressed the role of environmental and energy factors in their locational analysis. The OBERS model uses a 'shift share' trend model that may capture only aggregate past trends. The Curtis Harris Model uses a transportation cost variable to reflect the locational factors in the regional economy (Equation 4 of the Harris Model). Regional econometric models sidestep locational variables and use a 'top down' approach of depicting regional economy as a satellite system of the national model (Klein and Glickman 1977).

MREED represents an early attempt to incorporate in a regional economic model (in addition to the traditional market and transportation variables) the environmental and energy variables that appear to be emerging as significant in location of regional economic activity. The effects of the environmental (amenity) variables have been experienced in the last two decades (Berry and Neils 1969, Perloff 1969). The effects of rising energy prices while dramatic since the oil embargo of 1973 have been evident for a longer period -- about a decade (Miernyk 1976(a), Miernyk 1976(b)).

We hope to capture the effects of environmental and energy variables on regional location by estimating MREED's parameters over the last decade.

---

[1] Perloff and Wingo, _op_. _cit_., pp. 211-12.

The primary and manufacturing sector activities whose locations are determined in the above fashion in MREED, are viewed as export activities in each region determining the output levels in each region. Regional income is determined in MREED by an integration of a Keynesian system and economic base theory. Labor market and energy demand specification round off MREED's view of a region's economy. The detailed rationale of the model is explained next.

Table 5-1 lists the names and number of sectors used in the model. These 37 sectors are identified by the relevant SIC categories. It must be noted that the first year's work proposed here is at the 37 sector detail, to take advantage of directly available regional data. MREED can be estimated at a later phase at the 99 sector detail (at which INFORUM employment is detailed). The notation and the model are presented in the following pages.

The regional unit in the model is the BEA economic area, each of which is a labor market area including the place-of-work and place-of-residence of its labor force. The BEA areas -- 173 in all and covering the entire U.S. -- are economic rather than political or administrative units, permitting economic analysis in finer geographical detail than do states and in finer geographic and industrial detail than do the SMSA's.[2] The BEA's do not divide SMSA's or counties and are usually named after the largest SMSA. They range in size from area 14 (New York) with 18.2 million inhabitants to area 101 (Scottsbluff) with 104,000 people.

II. THE MODEL EQUATIONS

The model equations (Figure 5-2) are grouped into "blocks" representing the different classes of production sectors, regional expenditures, employment and energy. Special attention is paid to the demand for energy, since energy policy has come to occupy a central position in public policy. Energy as a composite good is consumed by households and commercial businesses as a final product and used by industry and transportation sectors as an intermediate input. Furthermore, energy is produced by combining the different sources of energy -- oil, gas, coal, electricity, solar energy. Demand for each of these sources will also be examined as part of the energy block of the model.

The major equations driving MREED are equations (1) and (3) that describe output by each of several primary and manufacturing industries. The output in each industrial sector in a region j is viewed as a function of input prices (W_{ijt}, EN_{jt}), size of the market (YA_{jt}) and regional environmental quality (EI_{vjt}). The variable YA_{jt} defines the size and spatial array of markets available to the industry located in region j at time t. It introduces space or the 'friction of distance' into the model by using a familiar form of operationalizing the concept of market potential (Stewart 1974, Lakshmanan

---

[2] See Regional Economic Analysis Division, "The BEA Economic Areas: Structural Changes and Economic Growth, 1950-73," Survey of Current Business, Nov. 1975, Vol. 55, No. 11, pp. 14-25.

TABLE 5-1. SECTORS PROPOSED FOR MREED PHASE I

	SIC NUMBERS
Primary	
Agriculture, forestry and fisheries	
Agriculture	01, 07
Forestry and fisheries	08, 09
Mining	
Metal	10
Coal	11, 12
Crude petroleum and natural gas	13
Nonmetallic, except fuels	14
Manufacturing	
Food and kindred products	20
Textile mill products	22
Apparel and other fabric products	23
Lumber products and furniture	24, 25
Paper and allied products	26
Printing and publishing	27
Chemicals and allied products	28
Petroleum refining	29
Primary metals	33
Fabricated metals	34, 19
Machinery, excluding electrical	35
Electrical machinery and supplies	36
Motor vehicles and equipment	371
Transportation equipment, excluding motor vehicles	37 except 371
Other manufacturing	21,30-32,38,39
Services	
Contract Construction	15-17
Transportation, communications and public utilities	
Railroad transportation	40
Trucking and warehousing	42
Other transportation and services	41,44,47
Communications	48
Utilities (electric, gas, sanitary)	49
Wholesale and retail trade	50,52-57,59
Finance, insurance and real estate	60-67
Other Services	
Lodging places and personal services	70, 72
Business and repair services	73, 75, 76
Amusement and recreation services	78, 79
Private households	88
Professional services	80,81,82,84,86,89
Government	
Civilian government	
Federal government	91 except Fed. military
State and local government	92, 93
Armed Forces	Part of 91

Source: BEA (US Dept. of Commerce) 1972 OBERS Projections of Regional Economic Activities in the U.S. Vol. 2, U.S.G.P.O. April 1974, p.1

Reproduced from best available copy.

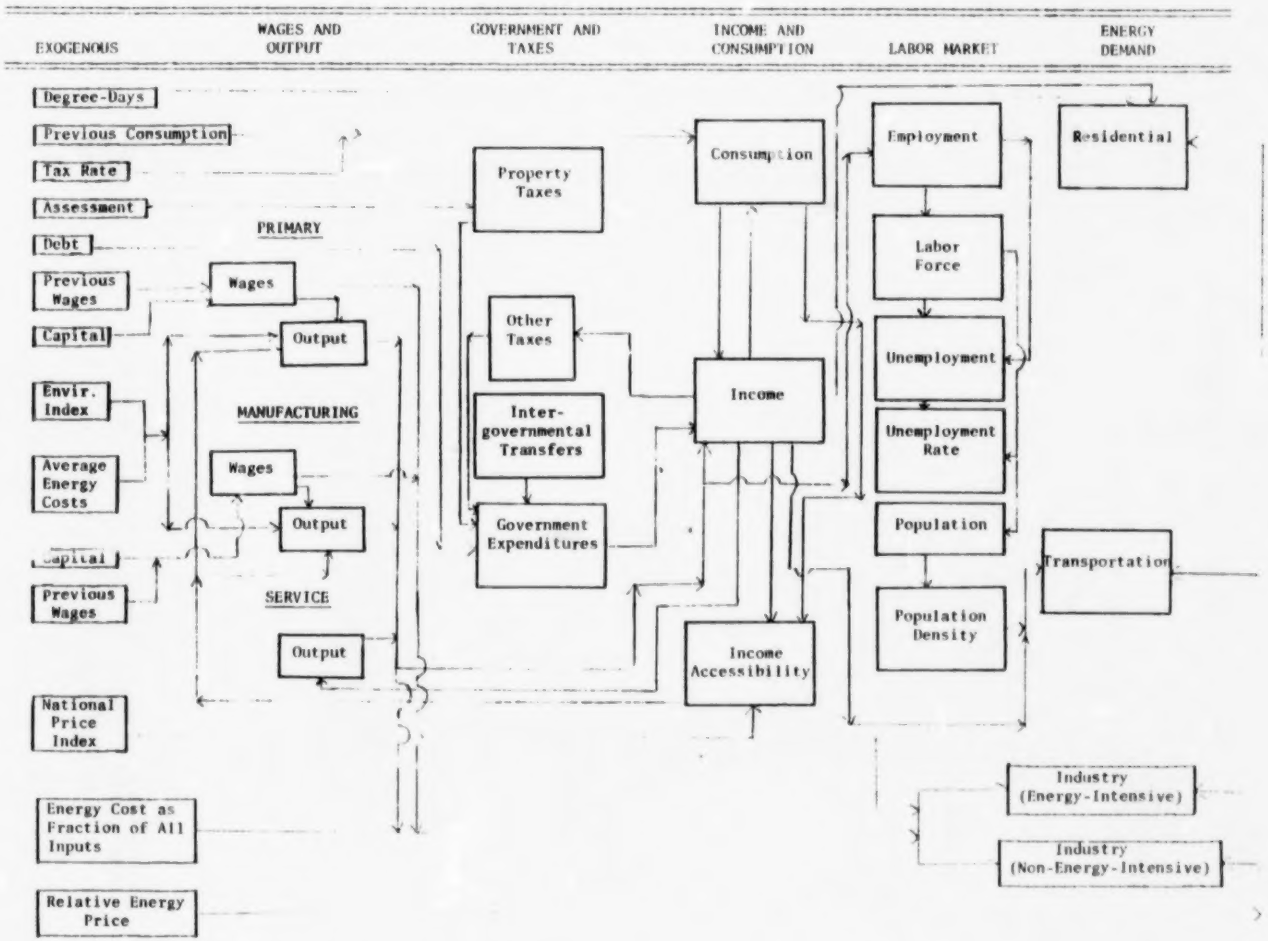


Figure 5-1. Structure of MREED Model.

Figure 5-2. The MREED Model.

Notation

AS_{jt} = Land assessment in region j in year t

C_{ij} = Generalized costs of transportation from i to j

C_{jt} = Personal consumption in region j in year t

CDD_j = Cooling degree days in region j

$Debt_{jt}$ = Total debt in region j in year t

EC_{kjt} = Energy input of k^{th} type of energy source into region j in year t

ED^{R}_{jt} = Energy demand in residential/commercial sector in region j in year t

ED^{II}_{jt} = Energy demand in energy intensive industrial sector in region j in year t

ED^{IE}_{jt} = Energy demand in non-energy intensive industrial sector in region j in year t

ED^{T}_{jt} = Energy demand in transportation sector in region j in year t

ED_{jt} = Energy demand in region j in year t

EI_{vjt} = Environmental index of type V in region j in year t

EN_{jt} = Average energy costs in region j in year t

G_{jt} = Government expenditures in region j in year t

HDD_j = Heating degree days in region j

K_{it} = Stock of capital in primary sector i in year t

LF_{jt} = Labor force in region j in year t

N^{P}_{ijt} = Employment in primary sector i in region j in year t

N^{M}_{ijt} = Employment in manufacturing sector i in region j in year t

(Continued)

Figure 5-2. (Continued)

N^S_{ijt} = Employment in service sector i in region j in year t

N_{it} = Employment in sector in in the region at t

N_{jt} = Total employment in region j at time t

P = National consumer price index

P^E = Relative price of energy

POP_{jt} = Population in region j in year t

$POPD_{jt}$ = Population densely in region j, in year t

RE_{it} = Ratio of cost of energy inputs to total costs in an industry

T_{jt} = Total tax receipts and transfer payments in region j in year t

T^P_{jt} = Total property tax receipts in region j in year t

T^O_{jt} = Total sales and income tax receipts in region j in year t

T^T_{jt} = Transfer payments in region j in year t

UN_{jt} = Unemployed in region j in year t

W^P_{ijt} = Wage rate in primary sector i in region j in year t

W^M_{ijk} = Wage rate in manufacturing sector i in region j in year t

W_{it} = National wage rate in sector i in year t

X^{II}_{jt} = Total output of all energy intensive industries in region j at time t

X^{IE}_{jt} = Total output of all non-energy intensive industries in region j at time t

X^P_{ijt} = Output of primary sector i in region j in year t

X^M_{ijt} = Output of manufacturing sector i in region j in year t

X^S_{ijt} = Output of service sector i in region j in year t

(Continued)

Figure 5-2. (Continued)

X_{jt} = Output of region j in year t

Y_{jt} = Income in region j at time t

Y^A_{jt} = Demand potential in (National income accessible to) region j at time t

Y = National income

r = Interest rate

u_{jt} = Unemployment rate in region j at time t

i = Industrial sector (i = 1,...,75)

j = Region (BEA economic area) (j = 1,...,173)

k = Type of energy source (k = 1,...,5)

t = Year

v = Type of environmental index

l_j = Average tax rate in the region

ϕ_j = Accessibility index for region j

(Continued)

Figure 5-2. (Continued)

THE MODEL

Output Block

Primary, Manufacturing, and Service Sectors

(1) $$X^P_{ijt} = f(Y^A_{jt}, EI_{vjt}, W^P_{ijt}, EN_{jt})$$

(2) $$W^P_{ijt} = f(W^P_{ijt-1}, K^P_{it}/N^P_{it})$$

(3) $$X^M_{ijt} = f(Y^A_{jt-1}, EI_{vjt}, W^M_{ijt}, EN_{jt})$$

(4) $$W^M_{ijt} = f(W^M_{ijt-1}, K^M_{it}/N^M_{it})$$

(5) $$Y^A_{jt} = \sum_j Y_{jt} \frac{e^{-\rho c_{ijt}}}{\phi_j}$$

(6) $$X^S_{ijt} = f(Y_{jt})$$

Income Block

(7) $$Y_{jt} = f(C_{jt}, \sum_i X_{ijt}, G_{jt})$$

(8) $$C_{jt} = f\{(1-1_j)Y_{jt}, C_{jt-1}\}$$

(9) $$G_{jt} = f(T_{jt}, Debt_{jt})$$

(10) $$T_{it} = T^P_{jt} + T^O_{jt} + T^T_{jt}$$

(11) $$T^P_{jt} = f(AS_{jt})$$

(12) $$T^O_{jt} = f(Y_{jt})$$

(13) $$T^T_{jt} = f\left(\frac{T^P_{jt-1} + T^O_{jt-1}}{T_{jt-1}}, \frac{T^P_{jt-1} + T^O_{jt-1}}{Pop_{jt-1}}, Y_{jt-1}\right)$$

(Continued)

Figure 5-2. (Continued)

Labor Market Block

$$(14) \quad N^{P}_{ijt} = f(X^{P}_{ijt}, t)$$

$$(15) \quad N^{M}_{ijt} = f(X^{M}_{ijt}, t)$$

$$(16) \quad N^{S}_{ijt} = f(X^{S}_{ijt}, t)$$

$$(17) \quad N_{jt} = \sum_{i} N_{ijt}$$

$$(18) \quad LF_{jt} = f(N_{jt}, t)$$

$$(19) \quad POP_{jt} = f(LF_{jt}, t)$$

$$(20) \quad UN_{jt} = LF_{jt} - N_{jt}$$

$$(21) \quad u_{jt} = UN_{jt}/LF_{jt}$$

Energy Block

Residential Energy Demand

$$(22) \quad ED^{R}_{jt} = f\{(1-1_{j})Y_{jt}, P^{E}_{j}, CDD_{j}, HDD_{j}\}$$

Industrial Energy Demand

$$(23) \quad ED^{II}_{jt} = f(X^{II}_{jt}, W_{it}/\rho, \rho^{E})$$

$$(24) \quad ED^{IE}_{jt} = f(X^{IE}_{jt}, W_{it}/\rho, \rho^{E})$$

(Continued)

Figure 5-2. (Continued)

Transportation Energy Demand

(25) $ED^{T}_{jt} = f(Y_{jt}, PQPD_{jt}, P^{E})$

(26) $ED_{jt} = ED^{R}_{jt} + ED^{I}_{jt} + ED^{T}_{jt}$

and Hansen 1965). Stated in equation (5), the demand potential or the size of the national market accessible to industries in region j is a weighted sum of demand in all regions j deflated by costs of interregional communication c_{ijt}.[3] The larger the regional income and the more aggregate proximity it has to other regions, the larger the markets accessible to it. The generalized communication costs c_{ijt}, include in theory money costs, time costs, and risk costs of transportation. If these costs are assumed to be proportional to distance, investments in transportation (e.g., Interstate System) reduce money and time costs. However, because of increased traffic, risk costs go up, effects will vary by sector, this will lead to larger market and supply potentials, increases in interregional trade and a trend towards larger national markets. On the other hand, increases in energy costs will cause increased transportation costs, with the effects most severe among those industries where transportation of goods is relatively important. Market potentials and interregional traffic will decline. Klassen has shown in such a case that smaller regions will lose more than larger regions since regional economies will move to some degree towards more autarky with less interregional trade and a stronger concentration on their own market (Klassen 1976). However the reduced traffic will also lead to lower congestion and air pollution, thus leading to improved environmental quality. The operational measure of Y^A_{jt} is described in the next section. In practice, a time lagged value of Y^A_{jt} will be used to facilitate ease of estimation.

The environmental quality variable used in equations (1) and (3) is intended not as one index but as a subset of indicators or measures that collectively convey information about the quality of the broad and complex environment in the regions. Some of the components of the environment such as pollutant concentration and congestion can be measured directly while many others cannot be (Thomas 1972, Hornback et al., 1974, Craik and Zube 1976). Several indicators may be integrated into one index for the more complex components where data are available (MITRE 1974, MITRE 1972, Liu 1975, CEQ 1972). However, the weighting involved in the development of indices is open to controversy (Thom and Ott 1975, Thomas 1972, Liu 1975). While we are aware of the conceptual and empirical difficulties, we hope to experiment and make a beginning with two or three key indicators. Those indicators will be selected that appear to affect regional location of firms and households and also respond to changes in environmental policy. Though it is difficult to specify these variables in any greater detail at this stage, we hope to use pollution indices and congestion indices for which data series can be developed.

---

[3] $\phi_j = \sum_j e^{-\rho c_{ijt}}$ = the reciprocal of the accessibility of region j to all other regions.

Regional wage rates and average energy costs -- the price variables of labor and energy inputs -- will be other independent variables. While national energy costs had been declining in the early and mid-sixties and have been since turning upward in the nation, regional variations have always been significant. We hypothesize that such regional variations will explain part of the regional differences in industrial location.

The wages (equations (2) and (4)) in the several primary and manufacturing sectors are viewed as functions of the lagged wages in the sector and capital/labor ratios. This form of wage equations was earlier estimated at the SMSA level by the principal investigator and others (Scully 1969, Lakshmanan, Lo and Krishnamoorthi 1972).

The regional output in the market oriented service sectors (equation 6) is determined by regional income. An integration of the Keynesian system and economic base theory appears in the income determination block of the model. The primary and manufacturing sectors are viewed as export oriented and their role in income determination is reflected by their combined output (equation (7)). Regional consumption expenditures C_{jt} and local government expenditures G_{jt} are other determinants of regional income in equation (7). Regional consumption is viewed as a function of disposable income and lagged consumption (equation (8)). Equations (9) through (13) describe the local government revenues and expenditures. Equation (9) is merely a simple relation between local government expenditure, G_{jt}, and total local government revenue, T_{jt}, and the public debt. Equation (10) is an identity that treats local government revenue as the sum of three components: local property tax, T^{P}_{jt}, other local taxes and revenues, T^{O}_{jt}, and federal and state transfer payment to the local government, T^{A}_{jt}. Equation (11) is a property tax equation that states that local property tax is a function of gross assessed property value, AS_{jt}. Equation (12) relates other local taxes and revenues to regional income Y_j. Finally, a behavior equation of federal and state transfer payment to local government is given in equation (13). Transfer payment to the local government, is determined by lagged values of total local revenue effort, per capita local tax, and regional income.

The equations predicting regional employment and population for a given year are provided in the labor market block. The equations determining employment in all sectors - primary, manufacturing, and service - are inverse forms of production function relating employment to output (equations (14), (15), and (16)). We follow Glickman in using a time variable which is intended to capture productivity changes (Glickman 1974). Equation (17) is an identity for total employment in the region. Labor force is related to the level of employment and a time variable (equation (18)). Labor force varies directly with employment opportunities. The effect of migration on the labor force is provided by the time variable t. Population is then related to labor force and the time variable which is intended to capture changes in the labor force participation rate (equation (19)). The equations for determining total unemployment and the unemployment rate are definitional (equations (20) and (21)).

The energy block describes the demand for energy by each major use group. The demand for residential energy use in a region is related to regional disposable income and the relative price of energy. In addition, the number of annual heating and cooling degree days in the region, which reflect the regional variations in space heating and cooling requirements are also used as independent variables (equation (22)). Industries are disaggregated into two groups -- energy intensive consumers (ED^{II}) and non-energy intensive consumers (ED^{IE}). The idea behind this distinction is that the energy consumption behavior of the former group in response to price changes is likely to be quite different from that of the latter group whose energy inputs are a smaller proportion of total inputs. Each group has a separate demand equation (in a later phase, we may estimate energy consumption by each of the manufacturing industries listed in Table 1, as was done by the principal investigator earlier -- Lakshmanan, Lo and Krishnamoorthi 1972). For both industrial groups, the demand for industrial energy is hypothesized to be related to output level, real wage and relative price of energy (equations (23) and (24)). The demand for transportation energy is a function of regional income and relative energy price. In addition, population density in the region is introduced into this equation as a proxy for level of use of public transportation (equation (25)). Equation (26) is an identity describing the total demand for energy.

III. DATA BASE

Data requirements for calibration of the model's equations pose some but tractable problems in that little data are presently available in fine detail by BEA economic area. The only substantial data publication has been the series of OBERS historical data and projected estimates for selected years.[4] However, county-level data may be aggregated up to BEA areas. Also, data by SMSA can likely be extrapolated to BEA areas, as the BEA classification essentially centers around one or more SMSA's and surrounding counties. Also, allocation of state figures to BEA areas may be permissible in some cases where reasonable assumptions can be presumed to hold.

A preliminary exploration of available data reveals standard series of federal government publications as the primary source of data. The economic censuses of the Bureau of the Census at five-year intervals provide county- and SMSA-level detail for many employment and production data requirements. The Labor Department's Bureau of Labor Statistics also provides detailed data on wages and employment, as well as price indices. Federal Power Commission publications provide energy data in great detail. The current intensive energy data gathering program in the Federal government will likely provide additional detailed data.

---

[4] In some cases, available BEA data may be used to scale unobserved data. For example, earnings by industry may be used to scale INFORUM sector output to BEA area level.

CHAPTER 6

A MODEL TO ASSESS REGIONAL ECONOMIC EFFECTS OF CHANGES IN FINAL DEMAND ON ENERGY PRICES

I. INTRODUCTION

This chapter seeks to present a model in which the effects of an exogenous change in one region's final demand or set of prices on interregional trade flows, regional production and value added can be analyzed.

Before the model is presented, there are a number of assumptions which need to be made clear. First, the model is a short-run model and therefore assumes the regional production coefficients and interregional trade coefficients are fixed. It is assumed that the period elapsing is short enough so as not to allow production and trade coefficients to respond to changes in input and output prices. In a longer-run, dynamic version of this model explicit account must be taken of substitution effects in production and consumption.

Initially, final demand is taken as exogenous. However, any changes in production will normally lead to a change in income, and hence a change in consumption and other elements of final demand. This aspect of income induced effects on final demand will be examined in the third section of this chapter. A review of some of the possible applications of this model will be taken up in the fourth and final section.

II. THE MODEL

In discussing the effect on interregional trade flows and regional production following an exogenous change in final demand or prices in one of the regions, one must realize that each region comprises a number of states which interact among themselves within the region and directly with the states in any other region. Taking this into consideration, we see that the two region model breaks down very rapidly to a complicated multiregional model. To avoid this problem, what follows is a presentation of a two region model, subsuming for the moment that each region consists of a number of smaller subregions.

Output Determination

$$(*)\quad \overset{X\,=}{\begin{bmatrix} X_1 \\ \hline X_2 \end{bmatrix}} = \overset{A^*}{\begin{array}{c@{}c} & \begin{array}{cc} \overbrace{\qquad}^{1} & \overbrace{\qquad}^{m} \end{array} \\ \begin{array}{c} 1\,\{ \\ m\,\{ \end{array} & \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right] \end{array}} \overset{X}{\begin{bmatrix} X_1 \\ \hline X_2 \end{bmatrix}} \overset{+}{+} \overset{F}{\begin{bmatrix} F_1 \\ \hline F_2 \end{bmatrix}}$$

where A_{11}^{1x1}, A_{21}^{1xm} refer to submatrices of coefficients showing the input of region 1's products to regions 1 and 2 respectively.

A_{21}^{mx1}, A_{22}^{mxm} refer to submatrices of coefficients showing the input of and 2 respectively.

X_1, X_2 are outputs in regions 1 and 2 respectively.

F_1, F_2 are final demands in regions 1 and 2 respectively.

The solution to this problem is:

$$X = \left[I - A^*\right]^{-1}F$$

where $\left[I-A^*\right]^{-1}$ is the standard Leontief inverse showing the total ultimate effects on production in the two regions following an exogenous change in F_1 or F_2.

Internal and External Multipliers --

Decomposition of the Leontief inverse using partitioned matrices yields the internal and external matrix multipliers for the two regions. The internal multiplier shows the inter-industrial propagation effects within each region, while the external multiplier shows the repercussion process in region 1 (2) due to the induced effects on output activities between the two regions.

Derivation of the Leontief Inverse --

By solving the following set of equations we can derive the Leontief inverse, which consists of the internal and external matrix multipliers for each region.

$$(1)\quad X_1 = A_{11}X_1 + A_{12}X_2 + F_1$$

$$(2)\quad X_2 = A_{21}X_1 + A_{22}X_2 + F_2$$

(a) Solve for X_1 and X_2 partially in each equation:

(3) $X_1 = [I - A_{11}]^{-1} A_{12} X_2 + [I - A_{11}]^{-1} F_1$

(4) $X_2 = [I - A_{22}]^{-1} A_{21} X_1 + [I - A_{22}]^{-1} F_2$

(b) Substitute (4) into (3).

(5) $X_1 = [I - A_{11}]^{-1} A_{12} \{[I - A_{22}]^{-1} A_{21} X_1 + [I - A_{22}]^{-1} F_2\} + [I - A_{11}]^{-1} F_1$

(5´) $X_1 = \{ I - [I - A_{11}]^{-1} A_{12} [I - A_{22}]^{-1} A_{21} \}^{-1} [I - A_{11}]^{-1} \{ F_1 + A_{12} [I - A_{22}]^{-1} F_2 \}$

The analogous expression for X_2 is:

(6) $X_2 = \{ I - [I - A_{22}]^{-1} A_{21} [I - A_{11}]^{-1} A_{12} \}^{-1} [I - A_{22}]^{-1} \{ A_{21} [I - A_{11}]^{-1} F_1 + F_2 \}$

To simplify notation, define the following:

$B_{11} = [I - A_{11}]^{-1}$

$B_{22} = [I - A_{22}]^{-1}$

$C^{2}_{11} = [I - B_{11} A_{12} B_{22} A_{21}]^{-1}$

$$C^{1}_{22} = [I - B_{22}A_{21}B_{11}A_{12}]^{-1}$$

Examining equation (3) or (4), we see that the partial derivative of X_1 (X_2) with respect to F_1 (F_2) yields B_{11} (B_{22}):

$$\frac{\partial X_1}{\partial F_1} = B_{11} = [I - A_{11}]^{-1}$$

$$\frac{\partial X_2}{\partial F_2} = B_{22} = [I - A_{22}]^{-1}$$

B_{11}, B_{22} can therefore be interpreted as the internal matrix multipliers for regions (1) and (2) respectively. Since each internal multiplier is not independent of the other region's industrial activity, but interregional propagations accompany the operation of the internal matrix multipliers, C^{2}_{11} and C^{1}_{22} can be interpreted as the external multipliers for regions (1) and (2) respectively.

In matrix notation equations (5′) and (6) become

$$(7)\quad \left[\begin{array}{c} X_1 \\ \hline X_2 \end{array}\right] = \left[\begin{array}{c|c} C^{2}_{11}B_{11} & C^{2}_{11}B_{11}A_{12}B_{22} \\ \hline C^{1}_{22}B_{22}A_{21}B_{11} & C^{1}_{22}B_{22} \end{array}\right] \cdot \left[\begin{array}{c} F_1 \\ \hline F_2 \end{array}\right]$$

where the diagonal term ($C^{j}_{ii}B_{ii}$) shows the ultimate total effect of repercussions within each region separately of an exogenous change in its own final demand, (F_i).

$$\text{eg:}\quad C^{j}_{ii}B_{ii} = [I - B_{ii}A_{ij}B_{jj}A_{ji}]^{-1}\,[I - A_{ii}]^{-1}$$

$$= \{[I - A_{ii}]\,[I - B_{ii}A_{ij}B_{jj}A_{ji}]\}^{-1}$$

$$= \{I - A_{ii} - B_{ii}A_{ij}B_{jj}A_{ji} + A_{ii}B_{ii}A_{ij}B_{jj}A_{ji}\}^{-1}$$

$$= \{ I - A_{ii} - A_{ij} B_{jj} A_{ji} [B_{ii} (I - A_{ii})]\}^{-1}$$

$$C^{j}_{ii} B_{ii} = \{ I - (A_{ii} + A_{ij} B_{jj} A_{ji}) \}^{-1}$$

The term in parentheses, $A_{ii} + A_{ij} B_{jj} A_{ji}$, is what Yamada and Ihara (1967) call the "Augmented Input Coefficient." It reflects the fact that production activity in region i, to begin with, needs its own goods-input (A_{ii}). At the same time, it also needs a certain amount of the jth region's goods as inputs (A_{ji}) to support its production. This supply for region i from region j induces the effect of production activity on region j (B_{jj}). To realize this production activity, region j in turn needs a certain amount of region i goods-input (A_{ij}).

The off-diagonal term in (7) ($C^{j}_{ii} B_{ii} A_{ij} B_{jj}$), by analogy, shows the ultimate total effect of repercussions within region i of an exogenous change in final demand of region j. For example, given a change in F_j, there will be set off a process of production within region j, reflected in B_{jj}. This will give rise to a demand for import-inputs from region i (A_{ij}). This, in turn, will induce production in i (B_{ii}), which when premultiplied by C^{j}_{ii} yields the ultimate total effect of repercussions in region i caused by a change in the demand for its exports.

Price Determination

The set of price formation equations pursuant to the output formation equations is given by

$$(8) \quad P_1 = A^T_{11} P_1 + A^T_{21} P_2 + V_1$$

$$(9) \quad P_2 = A^T_{12} P_1 + A^T_{22} P_2 + V_2$$

where V_1 and V_2 are vectors of value added in regions 1 and 2 respectively. Along the same lines used to solve for X_1 and X_2 we can solve for P_1 and P_2 as functions of the value added in each region.

$$(10) \quad P_1 = [I - A^T_{11}]^{-1} A^T_{21} P_2 + [I - A^T_{11}]^{-1} V_1$$

$$(11)\quad P_2 = [I - A_{22}^T]^{-1} A_{12}^T P_1 + [I - A_{22}^T]^{-1} V_2$$

$$(12)\quad P_1 = \{I - [I - A_{11}^T]^{-1} A_{21}^T [I - A_{22}^T]^{-1} A_{12}^T\}^{-1}$$

$$[I - A_{11}^T]^{-1} \{V_1 + A_{21}^T [I - A_{22}^T] V_2\}$$

$$(13)\quad P_2 = \{I - [I - A_{22}^T]^{-1} A_{12}^T [I - A_{11}^T]^{-1} A_{21}^T\}^{-1}$$

$$[I - A_{22}^T]^{-1} \quad \{A_{12}^T [I - A_{11}^T] V_1 + V_2\}$$

Letting

$B_{ii}^T = [I - A_{ii}^T]^{-1}$ $\quad i = 1,2$ $\quad$ T≡transpose

$C^{2T}_{11} = [I - B^T_{11} A^T_{21} B^T_{22} A^T_{12}]^{-1}$

$C^{1T}_{22} = [I - B^T_{22} A^T_{12} B^T_{11} A^T_{21}]^{-1}$

we can rewrite

$$\left[\begin{array}{c} P_1 \\ \hline P_2 \end{array}\right] = \left[\begin{array}{c|c} C^{2T}_{11} B^T_{11} & C^{2T}_{11} B^T_{22} \quad A^T_{21} B^T_{22} \\ \hline C^{1}_{22} B^T_{22} A^T_{12} B^T_{11} & C^T_{22} B^T_{22} \end{array}\right] \cdot \left[\begin{array}{c} V_1 \\ \hline V_2 \end{array}\right]$$

$$P = [I - D^*]^{-1} V$$

or

$$[I - D^*] P = V$$

where the relationship between prices and value added in the two regions is analogous to the relationship between regional outputs and final demand. From the perspective of price determination, the problem is to determine the effect on value added of a change in any element of the price vector in the two regions.

III. INCOME INDUCED EXPENDITURES

In the Leontief 'Open End' system, final demand is treated exogenously. In a theoretical model of a closed economy, however, consumption and investment expenditures are hypothesized to be functions of income, or value added. This section attempts to incorporate the income induced effects of consumption and investment expenditure on output, and thus close the Leontief system.

The introduction of the income multiplier process into the Leontief system is largely the work of Miyazawa and Masegi (1963). We will first present the model in a national or uni-regional setting, then, we will present the model in the two region case with a revised interpretation from that given in Section II.

Uni-regional case

In the model presented in this paper consumption and investment expenditures, domestic absorption, are assumed to be a linear function of the level of income. Government expenditures and net exports, on the other hand, are assumed to compose the autonomous part of final demand.

The equation determining output can, therefore, be written as follows:

$$(1) \quad X = AX + Q + Z$$

where: Q is domestic absorption = $C + I$, a $(n \times 1)$ column vector.
Z is government expenditures plus net exports a $(n \times 1)$ column vector.
$Q + Z = F$ = final demand
X is a column vector $(n \times 1)$ of final outputs
A is a $(n \times n)$ matrix of technical coefficients.

Since domestic absorption is a linear function of income, Q can be written as follows:

$$(2) \quad Q = EVX$$

$$= \sum_{ij}\sum e_i v_j x_j$$

where
- E is a $(n \times n)$ column vector of coefficient of domestic absorption. For the i^{th} element, $e_j = E_i/Y$ is the expenditure share in total income of the i^{th} commodity
- V is a $1 \times n$ row vector of value added ratios. For the j^{th} element, $v_j = Y_j/x_j$ is the income earned in industry j.

X is a (n x 1) column vector of outputs.

Substituting (2) into (1) we obtain:

$$(3)\quad X = AX + EVX + Z$$

Solving (3) for x yields:

$$(4)\quad X = I - A - EV \quad {}^{-1} Z$$

An alternative way of writing (4) is:

$$(5)\quad X = (I-A)^{-1}\left[I-EV(I-A)^{-1}\right]^{-1}Z$$

where $\left[I-EV(I-A)^{-1}\right]^{-1}$ is what Miyazawa and Masegi refer to as the "Subjoined Inverse Matrix (1963, p. 92)." Formula (5) distinguishes the inverse reflecting expenditure activity from the inverse reflecting production activity.[1]

Two-Region Case

In the two-region case we must separate the income induced from the autonomous components of final demand for both regions. The set of equations for the two-region case is:

$$(6)\quad X_1 = A_{11}X_1 + A_{12}X_2 + E_1V_1X_1 + Z_1$$

$$X_2 = A_{21}X_1 + A_{22}X_2 + E_2V_2X_2 + Z_2$$

where . A_{11}, A_{21} are respectively 1 x 1 and 1 x m coefficient matrices showing the input of region 1's products to regions 1 and 2

. A_{12}, A_{22} are respectively m x 1 and m x m coefficient matrices showing the input of region 2's products to regions 1 and 2.

Solving the set of equations (6) for X_1 and X_2 yields:

$$(7)\quad X_1 = \{ I - [I-A_{11} - E_1V_1]^{-1} A_{12} [I-A_{22} - E_2V_2]^{-1} A_{21} \}^{-1}$$

$$\cdot\ [I-A_{11} - E_1V_1]^{-1} \{Z_1 + A_{12} \left[I - A_{22} - E_2V_2\right]^{-1} Z_2\}$$

---

[1] Note, the expenditure coefficients and value added ratios are assumed to be fixed during the period of analysis. Over time, however, they as the a_{ij}'s are subject to change.

$$X_2 = \{I - \left[I-A_{22} - E_2 V_2\right]^{-1} A_{21} \left[I-A_{11} - E_1 V_1\right]^{-1} A_{12}\}^{-1}$$

$$\cdot \left[I-A_{22} - E_2 V_2\right]^{-1} \{A_{21} \left[I-A_{11} - E_1 V_1\right]^{-1} Z_1 + Z_2\}$$

To simplify notation, define the following:

$$\tilde{B}_{ii} = [I-A_{ii} - E_i V_i]^{-1} \qquad i = 1,2$$

$$\tilde{C}^{2}_{11} = [I-B_{11} A_{12} B_{22} A_{21}]^{-1}$$

$$\tilde{C}^{1}_{22} = [I-B_{22} A_{21} B_{11} A_{12}]^{-1}$$

where . $\tilde{B}_{ii}$ is the internal matrix multiplier for each region incorporating the effect of income induced expenditures in the production process.

. $\tilde{C}^{j}_{ii}$ is the external multiplier for each region incorporating the influence of income on each region's expenditures and output and interregional trade flows.

In matrix notation (7) becomes:

$$(8) \quad \left[\begin{array}{c} X_1 \\ \hline X_2 \end{array}\right] = \left[\begin{array}{c|c} \tilde{C}^{2}_{11} \tilde{B}_{11} & \tilde{C}^{2}_{11} \tilde{B}_{11} A_{12} \tilde{B}_{22} \\ \hline \tilde{C}^{1}_{22} \tilde{B}_{22} A_{21} \tilde{B}_{11} & \tilde{C}^{1}_{22} \tilde{B}_{22} \end{array}\right] \cdot \left[\begin{array}{c} Z_1 \\ \hline Z_2 \end{array}\right]$$

IV. APPLICATIONS OF THE MODEL

As indicated in the previous sections of this chapter, this model can be used to analyze the effects on regional output, income, private expenditures, and interregional trade flows of an exogenous change in public expenditures in either of the two regions.

Furthermore, the effect of an exogenous change in a region's vector of prices will first be felt in that region's level of income, and will influence output, private expenditures and interregional trade flows through the income multiplier process.

This model, however, can also be used in an entirely different way. The two regions can be viewed as two distinct industrial sectors - industrial goods and agricultural goods producing sectors, energy-intensive and non-energy intensive sectors, etc. As with the analysis in the two region case, the effect on output, income, prices and intersectoral flows can be analyzed. By categorizing the industries in a national I-O table into energy producing and non-energy producing industries, for example, we can examine the degree of dependence (independence) of non-energy producing industries on the energy sector. This can be done by examining how non-energy output and intersectoral trade flows respond to an output expansion in the energy sector. Assuming, however, fixed production and trade coefficients, one must realize the dependence of one sector on the other will be overstated.

Similarly, we can analyze the cost-push inflation caused by increased energy prices. Again by dividing the national I-O table of prices and wages into those pertaining to energy and non-energy producing sectors the effect of energy price increases on other prices can be determined. However, we must again realize that by ignoring the price sensitivity of inputs, the extent of 'cost-push' inflation caused by energy price increases will be overstated.

V. CONCLUSION

As long as we are willing to accept the limitations of input-output analysis and have the resources to collect national and regional data, the model developed by Miyazawa and generalized by Yamada and Ihara will prove to be an invaluable instrument to analyze national, regional, and interregional economic repercussions of alternative energy and public expenditure policies.

CHAPTER 7

APPENDIX--DATA SOURCES FOR CHAPTER 3

I. EMPLOYMENT DATA

Tables 7-1 and 7-2 indicate the sources of employment data used in the study of the use of location quotients to regionalize the national input-output table (Chapter 3).

TABLE 7-1. SOURCES USED FOR EMPLOYMENT DATA FOR INDUSTRY SECTORS

AS = *Agricultural Statistics 1969*, U.S. Department of Agriculture, U.S.G.P.O., Washington, 1969.

BLS-SA = Bureau of Labor Statistics, U.S. Department of Labor, *Employment and Earnings, States and Areas, 1934-68*, Bull. 1370-6, 1969.

BLS-US = Bureau of Labor Statistics, U.S. Department of Labor, *Employment and Earnings, United States, 1909-75*, Bull. 1312-10, 1976.

CA = Social and Economic Statistics Administration, Bureau of the Census, U.S. Department of Commerce, *1969 Census of Agriculture*, U.S.G.P.O., Washington, 1973.

CBP = Bureau of the Census, U.S. Department of Commerce, *County Business Patterns, 1967*, U.S.G.P.O., Washington, 1968.

CG = Bureau of the Census, U.S. Department of Commerce, *1967 Census of Governments*, U.S.G.P.O., Washington, 1969.

CM = Bureau of the Census, U.S. Department of Commerce, *1967 Census of Manufactures*, U.S.G.P.O., Washington, 1971.

HFW = Wage and Hour and Public Contracts Division, U.S. Department of Labor, *Hired Farmworkers: A Study of the Effects of the $1.00 Minimum Wage Under the Fair Labor Standards Act, Submitted to the Congress, 1968*, January 1968, Table 1.

MY = Bureau of Mines, U.S. Department of Interior, *Minerals Yearbook, 1967*, U.S.G.P.O., Washington, 1968.

USN = Office of Civilian Personnel, U.S. Navy.

TABLE 7-2. SOURCES USED FOR EMPLOYMENT DATA, BY INDUSTRY SECTOR[a]

SECTOR	WASHINGTON	UNITED STATES
1. Field Crops[b]	HFW	HFW
2. Vegetables & Fruit[b]	CA, v.2, ch.4, p.442	AS, p.58
3. Livestock[b]	AS, p.58	
4. Fish, Forestry, & Agriculture Services	CBP, 49, p.3 CG, 7, p.47, p.22 AS, p.58 CA, v.2, ch.4, p.442 HFW	CBP, 1, p.6 CG, 3, p.20 AS, p.58 CA, p.442, v.2, ch.4 HFW
5. Meat Products	CBP, 49, p.4	CBP, 1, p.7
6. Dairy Products	CBP, 49, p.4	CBP, 1, p.7
7. Canning & Preserving	CBP, 49, p.4	CBP, 1, p.7
8. Grain Mills	CBP, 49, p.4	CBP, 1, p.7
9. Beverages	CBP, 49, p.4	CBP, 1, p.7
10. Other Foods	CM, III, p.48-7	CM, I, p.28[c]
11. Textiles	CBP, 49, p.5	CBP, 1, p.8
12. Apparel	CBP, 49, p.5	CBP, 1, p.8
13. Mining	CBP, 49, p.4	CBP, 1, p.6
14. Logging	CBP, 49, p.5	CBP, 1, p.9
15. Sawmills	CBP, 49, p.5	CBP, 1, p.9
16. Plywood	CBP, 49, p.5	CBP, 1, p.9
17. Other Wood	CBP, 49, p.5	CBP, 1, p.9
18. Furniture & Fixtures	CBP, 49, p.5	CBP, 1, p.9
19. Pulp Mills	CBP, 49, p.5	CBP, 1, p.9
20. Paper Mills	CBP, 49, p.5	CBP, 1, p.9
21. Paperboard Mills	CM, III, p.48-8	CM, 1, p.32[d]

(cont'd)

TABLE 7-2. SOURCES USED FOR EMPLOYMENT DATA, BY INDUSTRY SECTOR (cont'd)

SECTOR	WASHINGTON	UNITED STATES
22. Printing, Publishing	CBP, 49, p.6	CBP, 1, p.10
23. Industrial Chemicals	CBP, 49, p.6	CBP, 1, p.10
24. Other Chemicals	CBP, 49, p.6	CBP, 1, p.10
25. Petroleum Refining	CBP, 49, p.6	CBP, 1, p.10
26. Glass and Stone	CM, III, p. 48-8	CM, I, pp.34,36
27. Cement and Clay	CBP, 49, p.8	CBP,1, p.11
28. Iron and Steel	CBP, 49, p.6	CBP,1, p.11
29. Nonferrous Metals	CM, III, pp.48-8,48-9	CM, I, p.36[e]
30. Aluminum	MY, III, p.834	CBP, 1, p.11[f]
31. Heavy Metal	CBP, 49, p.7	CBP, 1, p.12
32. Light Metal	CBP, 49, p.7	CBP, 1, p.12
33. Nonelectric Equip.	CBP, 49, p.7	CBP, 1, p.12
34. Machine Tools	CBP, 49, p.7	CBP, 1, pp.12-13
35. Industrial Equipment	CBP, 49, p.7	CBP, 1, pp.12-13
36. Electric Machinery	CBP, 49, p.7	CBP, 1, p.13
37. Aerospace	CBP, 49, p.7	CBP, 1, p.13
38. Motor Vehicles	CBP, 49, p.7	CBP, 1, p.13
39. Ship Building	CBP, 49, p.7 USN	CBP, 1, p.13 USN
40. Other Manufacturing	CM, III, pp.48-7,48-8,48-10[g] CBP, 49, pp.6,8	CBP, 1, pp.7,8,10,13,14
41. Transportation	BLS-SA, pp.538-9	BLS-US, pp.592-601[h]
42. Utilities	CBP, 49, p.8 CG, v.7, 47, p.22	CBP, 1, p.15[i] CG, v.3, p.21
43. Communication	CBP, 49, p.8	CBP, 1, p.15

(cont'd)

TABLE 7-2. SOURCES USED FOR EMPLOYMENT DATA, BY INDUSTRY SECTOR (cont'd)

SECTOR	WASHINGTON	UNITED STATES
44. Construction	CBP, 49, pp. 4,10	CBP, 1, pp. 7,17
45. All Trade	CBP, 49, pp. 8-9 CG, v.3, p.78	CBP, 1, pp. 15-16 CG, v.3, p 20
46. Finance	CBP, 49, pp. 10-11	CBP, 1, p. 17
47. Insurance	CBP, 49, p. 10	CBP, 1, p. 17
48. Real Estate	CBP, 49, p. 10	CBP, 1, p. 17
49. Business Services	CBP, 49, pp. 11-12	CBP, 1, pp. 18-19
50. Personal Services	CBP, 49, pp. 11-12 CG, v.7, 47, p.22	CBP, 1, pp. 17-19[j]

Notes:

[a] The full titles of the data sources are abbreviated as in Table 7-1.

[b] Agricultural employment by sector was estimated for 1967 rather than combined, in order to minimize aggregation of the table. For Washington, estimates were calculated by crop, as the state's 1969 fraction of U.S. hired workers for over 150 days (CA), times U.S. 1967 hired employment for that crop (HFW), times the ratio of total 1967 Washington farm workers to hired 1967 Washington farm workers (AS). The totals for each crop in a sector were then normalized so that their total sum equalled the 1967 total. U.S. employment was derived, by crop, by multiplying national 1967 hired workers for that crop (HFW) by the 1967 national ratio of total to hired farm workers (AS). Again these figures were summed over all crops in a sector, then normalized.

[c] CM was used because of CBP disclosure problems.

[d] CM was used because of CBP disclosure problems.

[e] CM was used because of CBP disclosure problems. For Washington CM also has disclosure problems, so the mean value of the upper and lower limits of the range was taken.

---

[f] CM was used because of CBP disclosure problems.

[g] Because of disclosure problems, upper bound of range was used for SIC 31, and lower bound was used for SIC 19.

[h] BLS was used in order to include railroad employment. State and local government public transit services, airports, port operations, and highways were included. Federal government transportation could not be included because it is not separable in the national input-output table.

[i] CG provided data on state and local government utilities. Federal electric utilities were also included, with data obtained from Tennessee Valley Authority, Southeast, Southwest, and Alaska Power Administrations of the Department of Interior, and 1970 Bonneville Power Authority (BPA) figures for Washington and all employers. The 1967 BPA figures were estimated by multiplying the 1970 figures by the ratio of 1967 to 1970 operation and maintenance costs (with the 1970 figure discounted by a private, non-farm industry productivity index).

[j] CG provided data on state and local government services (health and hospitals, local libraries, parks and recreation).

II. POTENTIAL SOURCES OF REGIONAL GROSS OUTPUT ESTIMATES BY INDUSTRY SECTOR

The 1963 Washington State input-output report (Philip J. Bourque, *et al.*, *The Washington Economy: An Input-Output Study*, University of Washington Graduate School of Business Administration and Washington State Department of Commerce and Economic Development, 1967) provides a reference of data sources used to estimate gross output in its 1963 table. Manufacturing sector data relies primarily on the Census Bureau's *Census of Manufactures*. Its *Census of Business* is also used. Various other state, federal, and trade publications are cited for the other sectors, and similar documents are likely available for most regions.

Similarly, *Definitions and Conventions of the 1967 Input-Output Study* (Social and Economic Statistics Administration, Bureau of Economic Analysis, Interindustry Economic Division, U.S. Department of Commerce, October 1974) provides a list of publicly available data sources for the national level, many of which contain regional data. The Census Bureau's *Census of Agriculture* is highly disaggregated and detailed by region. Trade publications and the Interior Department's annual *Minerals Yearbook* also provide output data.

REFERENCES

CHAPTER 1.

1. Hitch, Charles J. (ed.), Modeling Energy-Economy Interactions: Five Approaches, Resources for the Future, Washington, D.C., 1977.

2. Hogan, William W. and S. Parikh, Comparison of Models of Energy and the Economy, Stanford Energy Modelling Forum Working Paper EMF 1, 4, February 1977.

3. House, Peter, Trading Off Environment, Economics and Energy, Lexington, Mass., D.C. Heath, 1977.

4. Lakshmanan, T.R. and S. Krishnamoorthi, The SEAS Test Model: Design and Implementation, U.S. Environmental Protection Agency, Washington, D.C., 1973.

CHAPTER 3.

1. Consad Research Corporation, Regional Federal Procurement Study, Washington, D.C., Office of Economic Research, U.S. Department of Commerce Contract 7-35211, October 1967.

2. Czamanski, Stanislaw and Emil E. Malizia, "Applicability and Limitations in the Use of National Input-Output Tables for Regional Studies," Papers of the Regional Science Association, xxiii, 1969, 65-78.

3. Goldman, Morris R., "Comments on Czamanski and Malizia," Papers of the Regional Science Association, xxiii, 1969, 79-80.

4. Isard, Walter, Methods of Regional Analysis, MIT, 1960.

5. Malizia, Emil and Daniel L. Bond, "Empirical Tests of the RAS Method of Interindustry Coefficient Adjustment," Journal of Regional Science, 14, no. 3, 1974, 355-65.

6. McMenamin, David G. and Joseph E. Haring, "An Appraisal of Nonsurvey Techniques for Estimating Regional Input-Output Models," Journal of Regional Science, 14, no. 2, 1974, 191-205.

7. Miernyk, William, "Comments of Czamanski and Malizia," Papers of the Regional Science Association, xxiii, 1969, 81-82.

8. Morrison, W.I. and P. Smith, "Nonsurvey Input-Output Techniques at the Small Area Level: An Evaluation," *Journal of Regional Science,* 14, no. 1, 1974, 1-14.

9. Round, J. I., "Regional Input-Output Models in the U.K.: A Re-Appraisal of Some Techniques," *Regional Studies,* 6, 1972, 1-9.

10. Schaffer, William and Kong Chu, "Nonsurvey Techniques for Constructing Regional Interindustry Models," *Papers of the Regional Science Association*, xxiii, 1969, 83-101.

11. Walderhaug, Albert, "State Input-Output Tables Derived from National Data," *Proceedings, American Statistical Association*, Business and Economic Statistics Section, 1971.

CHAPTER 4.

1. Crow, Robert T. "A Nationally-Linked Regional Econometric Model," *Journal of Regional Science*. Vol. 13, No. 2, 1973, 187-204.

2. Glickman, Norman J., "Son of the Specification of Regional Econometric Models," *Papers of the Regional Science Association*, Vol. 32, 1974.

3. Glickman, Norman J., "An Econometric Forecasting Model for the Philadelphia Region," *Journal of Regional Science*, Vol. 11, No. 1, 1971.

4. Harris, Curtis, *Urban Economies, 1985: A Multiregional, Multi-Industry Forecasting Model*, Lexington Books, (D. C. Heath), Lexington, Mass., 1973.

5. Klein, Lawrence R., "The Specification of Regional Econometric Models," *Papers of the Regional Science Association*, Vol. 23, 1969.

6. Klein, Lawrence R., and Norman J. Glickman, "Econometric Model Building at Regional Level," *Regional Science and Urban Economics*, 1977, 3-23.

7. Matilla, John M., "A Metropolitan Income Determination Model and the Estimation of Metropolitan Income Multipliers," *Journal of Regional Science*, Vol. 13, No. 1, 1973.

8. Miernyk, William H., "Regional and Interregional Input-Output Models: A Reappraisal," in Perlman, Leven and Chinitz (eds.), *Spatial, Regional and Population Economics*, Gordon and Breach, Science Publishers, Ltd., 1973.

9. Moses, Leon N., "A General Equilibrium Model of Production, Interregional Trade and Location of Industry," *Review of Economics and Statistics*, 1966, 42, 373-397.

10. Olsen, R. J., "An Overview of MULTIREGION," A paper presented at Seventh Annual Conference on Modeling and Simulation, Pittsburgh, PA, April 1976.

11. Polenske, Karen R., State Estimates of Technology, 1963, D.C. Heath and Company, Lexington, Mass., 1973.

12. Polenske, Karen R., A Guide for Users of the U.S. Multiregional Input-Output Model (revised version), PB-242-558, Department of Commerce, March 1974.

13. U. S. Water Resources Council, Concepts, Methodology, and Summary Data, Vol. I of the 1972 OBERS Projections, Economic Activity in the U.S., April 1974.

CHAPTER 5.

1. Adams, F. G., C. Brooking and N. J. Glickman, "On the Specification and Simulation of a Regional Econometric Model: A Model of Mississippi," Discussion Paper No. 266, Department of Economics, University of Pennsylvania, Philadelphia, 1973.

2. Berry, Brian J. L. and Elaine Neils, "Location Size and Shape of Cities as Influenced by Environmental Factors: The Urban Environment Unit Large," in The Quality of the Urban Environment. (Ed.) H. S. Perloff, Johns Hopkins University Press, Baltimore, 1969, 257-302.

3. (CEQ) Council on Environmental Quality, The Third Annual Report of the Council on Environmental Quality, U.S. Government Printing Office, Washington, D.C. 1972.

4. Chinitz, B., "The Effect of Transportation Forms on Regional Economic Growth," Traffic Quarterly, 14, 1960.

5. Craik, K. H. and E. H. Zube, Perceiving Environmental Quality. Plenum Press, New York, 1976.

6. Crow, R. T., "A Nationally Linked Regional Econometric Model," Journal of Regional Science, XIII, 1973, 187-204.

7. Dutta, M., and V. Su, "An Econometric Model of Puerto Rico," The Review of Economic Studies, XXXVI, 1969, 319-333.

8. L'Esperance, W. L., G. Nestel and G. Fromm, "Gross State Product and an Econometric Model of a State," American Statistical Association Journal, LXIV, 1969, 787-807.

9. Giarratani, Frank and Charles F. Socher, "The Pattern of Industrial Location and Rising Energy Prices," Atlantic Economic Journal, V, 1, March 1977, 48-55.

10. Glickman, N. J., "Son of 'The Specification of Regional Econometric Models'," *Papers of the Regional Science Association*, 32, 1974.

11. Glickman, N. J., "An Econometric Model of the Philadelphia Region," Ph.D. dissertation, University of Pennsylvania, Philadelphia, 1969.

12. Glickman, N. J., "An Econometric Model of the Philadelphia Region," *Journal of Regional Science*, XI, 1971, 15-32.

13. Harris, Chauncey D., "The Market as a Factor in the Localization of Industry in the U.S.," *Annals of the Association of American Geographers*.

14. Harris, Curtis C., *The Urban Economies, 1985*. Lexington, Mass.: D. C. Heath and Co., 1973.

15. Hornsback, Kenneth E., J. Guttman, H. L. Himmalstein, A. Rappaport and R. Reyna, *Studies in Environment, Vol. II, Quality of Life*. Prepared for U. S. EPA. EPA-600/5-73-0126, Washington, D.C., U.S. GPO, February 1974.

16. Kendrick, J. W. and C. M. Jaycox, "The Concept and Estimation of Gross State Product," *Southern Economic Journal*, XXXII, 1965, 153-168.

17. Klassen, Leo H., "Energy, Environment and Regional Development," in *Environment, Regional Science and Interregional Modelling*, (eds.) M. Chatterji and P. V. Rompuy. Springer-Verlag, Berlin, 1976, 1-15.

18. Lakshmanan, T. R. and Fu-chen Lo, "A Regional Economic Model for the Assessment of Effects of Air Pollution Abatement," *Environment and Planning*, Vol. 4, 1972, 73-97.

19. Lakshmanan, T. R., F. Lo and Krishnamoorthi, *The OAP Regional Econometric Model: A Revised Version*, Prepared for Environmental Protection Agency by CONSAD, NTIS Publication, Pittsburgh, September 1972.

20. Lakshmanan, T. R. and Walter G. Hansen, "A Retail Market Potential Model," *Journal of the American Institute of Planners*, May 1965.

21. Liu, Ben-chieh, *Quality of Life Indicators in the U.S. Metropolitan Areas, 1970*, Prepared for the U.S. Environmental Protection Agency, 1975.

22. Mattila, J. M., "A Metropolitan Income Determination Model and the Estimation of Metropolitan Income Multipliers," *Journal of Regional Science*, XIII, 1973, 1-16.

23. Miernyk, William H., "Some Regional Impacts of the Rising Costs of Energy," *Papers and Proceedings of the Regional Science Association*, 1976(b), 201-212.

24. Miernyk, William H., "Regional Economic Consequences of High Energy Prices in the U.S." *The Journal of Energy and Development*, Spring 1976, 213-239.

25. MITRE Corporation, *National Environmental Indices: Air Quality and Outdoor Recreation*, MTR 6159, April 1972.

26. MITRE Corporation, *The PDI Index*, Working Paper 7693, Washington, D.C., September 1974.

27. Perloff, Harvey S. and Loudon Wingo Jr., "National Resource Endowment and Regional Economic Growth," in J. J. Spengler (ed.), *Natural Resources and Economic Growth*, Conference Papers published by RFF, 1961, 191-212.

28. Perloff, Harvey S., "A Framework for Dealing with the Urban Environment," in *The Quality of the Urban Environment*, (ed.) H. S. Perloff, The Johns Hopkins University Press, Baltimore 1969, 3-25.

29. Polenske, K., *A Multiregional Input-Output Model for the United States*. Cambridge: Harvard Economic Research Project, 1970.

30. Scully, G. W., "Interstate Wage Differentials: A Cross-Section Analysis," *American Economic Review*, Vol. LIX, 1969, 757-773.

31. Stewart, J. Q., "Empirical Mathematical Rules Concerning the Distribution and Equilibrium of Population," *Geographic Review*, 37, 1947.

32. Thom, Gary C. and Wayne R. Ott, *Air Pollution Indices*, Prepared for CEQ and Environmental Protection Agency, December 1975.

33. Thomas, William A. (ed.), *Indicators of Environmental Quality*, Plenum Press, New York, 1972.

34. Ullman, Edward L., "Regional Development and the Geography of Concentration," *Papers and Proceedings of the Regional Science Association*, 4, 1958.

CHAPTER 6.

1. Polenske, Karen R., *State Estimates of Technology, 1963*, D. C. Heath and Co., 1974.

2. Polenske, Karen R., *A Guide for Users of the U.S. Multiregional Input-Output Model (revised version)*, PB-242-558, Department of Commerce, March 1974.

3. Miyazawa, K., "Internal and External Matrix Multipliers in the Input-Output Model," *Hitotsubashi Journal of Economics*, Vol. 7, No. 1, 1966, 38-55.

4. Miyazawa, K. and S. Masegi, "Interindustry Analyses and the Structure of Income Distribution," *Metroeconomica*, Vol. XV, Fas 2, 3 (1963), 89-103.

5. Yamada, H. and T. Ihara, "Input-Output Analysis of Interregional Repercussion," *Papers and Proceedings of the 3rd Far East Conference of the Regional Science Association*, Tokyo, Japan, 1967, Vol. III, 3-31.

*U S GOVERNMENT PRINTING OFFICE: 1980 311-132/70